

JS011312752B2

(12) United States Patent

Bowen et al.

(10) Patent No.: US 11,312,752 B2

(45) **Date of Patent:**

Apr. 26, 2022

(54) INSECT INHIBITORY PROTEINS

(71) Applicant: Monsanto Technology LLC, St. Louis, MO (US)

(72) Inventors: **David J. Bowen**, Wildwood, MO (US); **Catherine A. Chay**, Ballwin, MO (US); **Arlene R. Howe**, Clarkson Valley, MO (US); **Uma Kesanapalli**, Chesterfield,

MO (US)

(73) Assignee: Monsanto Technology LLC, St. Louis,

MO (US)

(*) Notice: Subject to any disclaimer, the term of this

patent is extended or adjusted under 35

U.S.C. 154(b) by 0 days.

(21) Appl. No.: 16/810,810

(22) Filed: Mar. 5, 2020

(65) Prior Publication Data

US 2020/0262876 A1 Aug. 20, 2020

Related U.S. Application Data

- (62) Division of application No. 15/727,883, filed on Oct.9, 2017, now Pat. No. 10,626,151.
- (60) Provisional application No. 62/406,082, filed on Oct. 10, 2016.

(51)	Int.	Cl.

C07K 14/32	(2006.01)
C12N 15/82	(2006.01)
C07K 14/325	(2006.01)
C12Q 1/6895	(2018.01)

(52) U.S. Cl.

CPC *C07K 14/325* (2013.01); *C12N 15/8286* (2013.01); *C12Q 1/6895* (2013.01)

(58) Field of Classification Search

None

See application file for complete search history.

(56) References Cited

U.S. PATENT DOCUMENTS

6,204,246 B	1 3/2001	Bosch et al.
6,780,408 B	1 8/2004	Bosch et al.
9,499,835 B	2 11/2016	Meade et al.
9,556,453 B	2 1/2017	Meade et al.
9,567,602 B	2 2/2017	Meade et al.
9,663,795 B	2 5/2017	Meade et al.
9,796,982 B	2 10/2017	Meade et al.

2003/0195336 A1 10/2003 Baum et al. 2004/0197916 A1 10/2004 Carozzi et al. 2004/0221334 A1 11/2004 Baum et al.

2008/0040827 A1* 2/2008 Donovan C07K 14/325 800/302

2009/0036377 A1 2/2009 Carozzi et al.

OTHER PUBLICATIONS

Honee et al., "Nucleotide sequence of crystal protein gene isolated from *B. thuringiensis* subspecies *entomocidus* 60.5 coding for a toxin highly active against *Spodoptera* species," *Nucleic Acids Res*, 16:6240, 1988.

Kao et al., "Cloning and expression of the insecticidal crystal protein gene Cry1Ca9 of *Bacillus thuringiensis* G10-01A from Taiwan granaries," *Curr Microbiol*, 47:295-299, 2003.

Palma et al., "Bacillus thuringiensis toxins: An overview of their biocidal activity," Toxins, 6:3296-3325, 2014.

Sanchis et al., "Nucleotide sequence and analysis of the N-terminal coding region of the *Spodoptera*-active delta-endotoxin gene of *Bacillus thuringiensis aizawai* 7.29," *Mol Microbiol*, 3:229-238, 1080

Smith et al., "Mosquitocidal activity of the CrylC d-endotoxin from *Bacillus thuringiensis* subsp. *aizawai," Appl Environ Microbiol*, 62:680-684, 1996.

Visser et al., "Genes from *Bacillus thuringiensis entomocidus* 60.5 coding for insect-specific crystal proteins," *Mol Gen Genet*, 212:219-224, 1988.

International Search Report and Written Opinion regarding International Application No. PCT/US2017/055731, dated Feb. 9, 2018.

* cited by examiner

Primary Examiner — Matthew R Keogh (74) Attorney, Agent, or Firm — Dentons US LLP; Timothy Ball

(57) ABSTRACT

Pesticidal proteins exhibiting toxic activity against Lepidopteran pest species are disclosed, and include, but are not limited to, TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL. DNA constructs are provided which contain a recombinant nucleic acid sequence encoding one or more of the disclosed pesticidal proteins. Transgenic plants, plant cells, seed, and plant parts resistant to Lepidopteran infestation are provided which contain recombinant nucleic acid sequences encoding the pesticidal proteins of the present invention. Methods for detecting the presence of the recombinant nucleic acid sequences or the proteins of the present invention in a biological sample, and methods of controlling Lepidopteran species pests using any of the TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL pesticidal proteins are also provided.

32 Claims, No Drawings

Specification includes a Sequence Listing.

INSECT INHIBITORY PROTEINS

REFERENCE TO RELATED APPLICATION

This application is a divisional of U.S. application Ser. ⁵ No. 15/727,883, filed Oct. 9, 2017, which application claims the benefit of U.S. provisional application No. 62/406,082, filed Oct. 10, 2016, each of which is herein incorporated by reference in its entirety.

INCORPORATION OF SEQUENCE LISTING

The file named "MONS426US_ST25.txt" containing a computer-readable form of the Sequence Listing was created on Oct. 8, 2017. This file is 77,030 bytes (measured in MS-Windows®), filed contemporaneously by electronic submission (using the United States Patent Office EFS-Web filing system), and incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

The invention generally relates to the field of insect inhibitory proteins. A novel class of proteins exhibiting 25 insect inhibitory activity against agriculturally-relevant pests of crop plants and seeds are disclosed. In particular, the disclosed class of proteins is insecticidally active against agriculturally-relevant pests of crop plants and seeds, particularly Lepidopteran species of insect pests. Plants, plant 30 parts, and seeds containing a recombinant polynucleotide construct encoding one or more of the disclosed toxin proteins are provided.

BACKGROUND OF THE INVENTION

Improving crop yield from agriculturally significant plants including, among others, corn, soybean, sugarcane, rice, wheat, vegetables, and cotton, has become increasingly important. In addition to the growing need for agricultural 40 products to feed, clothe and provide energy for a growing human population, climate-related effects and pressure from the growing population to use land other than for agricultural practices are predicted to reduce the amount of arable land available for farming. These factors have led to grim 45 forecasts of food security, particularly in the absence of major improvements in plant biotechnology and agronomic practices. In light of these pressures, environmentally sustainable improvements in technology, agricultural techniques, and pest management are vital tools to expand crop 50 production on the limited amount of arable land available for farming.

Insects, particularly insects within the order Lepidoptera and Coleoptera, are considered a major cause of damage to field crops, thereby decreasing crop yields over infested 55 areas. Lepidopteran pest species which negatively impact agriculture include, but are not limited to, Beet armyworm (Spodoptera exigua), Corn earworm (Helicoverpa zea), Cotton leaf worm (Alabama argillacea), European corn borer (Ostrinia nubilalis), Fall armyworm (Spodoptera frugiperda), Old World bollworm (Helicoverpa armigera), Oriental leaf worm (Spodoptera litura), Pink bollworm (Pectinophora gossypiella), Cry1Ac resistant Pink bollworm (Pectinophora gossypiella), Soybean looper (Chrysodeixis includens), Southern armyworm (Spodoptera eridania), Southwestern corn borer (Diatraea grandiosella), Spotted bollworm (Earias vittella), Sugarcane borer (Diatraea (Diatraea))

2

traea saccharalis), Tobacco budworm (Heliothis virescens), and Velvet bean caterpillar (Anticarsia gemmatalis).

Historically, the intensive application of synthetic chemical insecticides was relied upon as the pest control agent in agriculture. Concerns for the environment and human health, in addition to emerging resistance issues, stimulated the research and development of biological pesticides. This research effort led to the progressive discovery and use of various entomopathogenic microbial species, including bactoria.

The biological control paradigm shifted when the potential of entomopathogenic bacteria, especially bacteria belonging to the genus Bacillus, was discovered and developed as a biological pest control agent. Strains of the bacterium Bacillus thuringiensis (Bt) have been used as a source for pesticidal proteins since it was discovered that Bt strains show a high toxicity against specific insects. Bt strains are known to produce delta-endotoxins that are localized within parasporal crystalline inclusion bodies at 20 the onset of sporulation and during the stationary growth phase (e.g., Cry proteins), and are also known to produce secreted insecticidal protein. Upon ingestion by a susceptible insect, delta-endotoxins as well as secreted toxins exert their effects at the surface of the midgut epithelium, disrupting the cell membrane, leading to cell disruption and death. Genes encoding insecticidal proteins have also been identified in bacterial species other than Bt, including other Bacillus and a diversity of additional bacterial species, such as Brevibacillus laterosporus, Lysinibacillus sphaericus ("Ls" formerly known as Bacillus sphaericus) and Paenibacillus popilliae.

Crystalline and secreted soluble insecticidal toxins are highly specific for their hosts and have gained worldwide acceptance as alternatives to chemical insecticides. For example, insecticidal toxin proteins have been employed in various agricultural applications to protect agriculturally important plants from insect infestations, decrease the need for chemical pesticide applications, and increase yields. Insecticidal toxin proteins are used to control agriculturally-relevant pests of crop plants by mechanical methods, such as spraying to disperse microbial formulations containing various bacteria strains onto plant surfaces, and by using genetic transformation techniques to produce transgenic plants and seeds expressing insecticidal toxin protein.

The use of transgenic plants expressing insecticidal toxin proteins has been globally adapted. For example, in 2012, 26.1 million hectares were planted with transgenic crops expressing Bt toxins (James, C., Global Status of Commercialized Biotech/GM Crops: 2012. ISAAA Brief No. 44). The global use of transgenic insect-protected crops and the limited number of insecticidal toxin proteins used in these crops has created a selection pressure for existing insect alleles that impart resistance to the currently-utilized insecticidal proteins.

The development of resistance in target pests to insecticidal toxin proteins creates the continuing need for discovery and development of new forms of insecticidal toxin proteins that are useful for managing the increase in insect resistance to transgenic crops expressing insecticidal toxin proteins. New protein toxins with improved efficacy and which exhibit control over a broader spectrum of susceptible insect species will reduce the number of surviving insects which can develop resistance alleles. In addition, the use in one plant of two or more transgenic insecticidal toxin proteins toxic to the same insect pest and displaying different modes of action reduces the probability of resistance in any single target insect species.

Thus, the inventors disclose herein a novel protein toxin family from Bacillus thuringiensis, along with similar toxin proteins, variant proteins, and exemplary recombinant proteins that exhibit insecticidal activity against target Lepidopteran species, particularly against Beet armyworm (Spo- 5 doptera exigua), Corn earworm (Helicoverpa zea), Cotton leaf worm (Alabama argillacea), European corn borer (Ostrinia nubilalis), Fall armyworm (Spodoptera frugiperda), Old World bollworm (Helicoverpa armigera), Oriental leaf worm (Spodoptera litura), Pink bollworm (Pectinophora 10 gossypiella), Cry1Ac resistant Pink bollworm (Pectinophora gossypiella), Soybean looper (Chrysodeixis includens), Southern armyworm (Spodoptera eridania), Southwestern corn borer (Diatraea grandiosella), Spotted bollworm (Earias vittella), Sugarcane borer (Diatraea sac- 15 charalis), Tobacco budworm (Heliothis virescens), and Velvet bean caterpillar (Anticarsia gemmatalis).

SUMMARY OF THE INVENTION

Disclosed herein is a novel group of pesticidal proteins with insect inhibitory activity (toxin proteins), referred to herein as TIC4472, TIC1425, and TIC2613 belonging to the TIC4472 protein toxin class, which are shown to exhibit inhibitory activity against one or more pests of crop plants. 25 The TIC4472 protein and proteins in the TIC4472 protein toxin class can be used alone or in combination with other insecticidal proteins and toxic agents in formulations and in planta, thus providing alternatives to insecticidal proteins and insecticide chemistries currently in use in agricultural 30 systems.

In one embodiment, disclosed in this application is a recombinant nucleic acid molecule comprising a heterologous promoter fragment operably linked to a polynucleotide wherein (a) said pesticidal protein comprises the amino acid sequence of SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (b) said pesticidal protein comprises an amino acid sequence having: (i) at least 93%, or 95%, or 98%, or 99%, or about 100% amino acid sequence identity to SEQ ID 40 NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (ii) at least 73%, or 75%, or 80%, or 85%, or 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:8 or SEQ ID NO:10; or (c) said polynucleotide segment hybridizes to a polynucleotide having the nucleotide sequence of SEQ ID 45 NO:3, SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:9; or (d) said polynucleotide segment encoding a pesticidal protein or fragment thereof comprises a polynucleotide sequence having at least 65%, or 70%, or 75%, or 80%, or 85%, or 90%, or 95%, or 98%, or 99%, or about 50 100% sequence identity to the nucleotide sequence of SEQ ID NO:3, SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:9; or (e) said recombinant nucleic acid molecule is in operable linkage with a vector, and said vector is selected from the group consisting of a plasmid, phagemid, 55 bacmid, cosmid, and a bacterial or yeast artificial chromosome. The recombinant nucleic acid molecule can comprise a sequence that functions to express the pesticidal protein in a plant; or is expressed in a plant cell to produce a pesticidally effective amount of pesticidal protein.

In another embodiment of this application, host cells comprising a recombinant nucleic acid molecule of the application are provided, wherein the host cell is selected from the group consisting of a bacterial and a plant cell. Contemplated bacterial host cells include Agrobacterium, 65 Rhizobium, Bacillus, Brevibacillus, Escherichia, Pseudomonas, Klebsiella, Pantoec, and Erwinia. In certain embodi-

ments, said Bacillus species is Bacillus cereus or Bacillus thuringiensis, said Brevibacillus is Brevibacillus laterosperous, or Escherichia is Escherichia coli. Contemplated plant host cells include a dicotyledonous plant cell and a monocotyledonous plant cell. Contemplated plant cells further include an alfalfa, banana, barley, bean, broccoli, cabbage, brassica, carrot, cassava, castor, cauliflower, celery, chickpea, Chinese cabbage, citrus, coconut, coffee, corn, clover, cotton (Gossypium sp.), a cucurbit, cucumber, Douglas fir, eggplant, eucalyptus, flax, garlic, grape, hops, leek, lettuce, Loblolly pine, millets, melons, nut, oat, olive, onion, ornamental, palm, pasture grass, pea, peanut, pepper, pigeonpea, pine, potato, poplar, pumpkin, Radiata pine, radish, rapeseed, rice, rootstocks, rye, safflower, shrub, sorghum, Southern pine, soybean, spinach, squash, strawberry, sugar beet, sugarcane, sunflower, sweet corn, sweet gum, sweet potato, switchgrass, tea, tobacco, tomato, triticale, turf grass, watermelon, and wheat plant cell.

In another embodiment, the pesticidal protein exhibits activity against Lepidopteran insects, including Beet armyworm (Spodoptera exigua), Corn earworm (Helicoverpa zea), Cotton leaf worm (Alabama argillacea), European corn borer (Ostrinia nubilalis), Fall armyworm (Spodoptera frugiperda), Old World bollworm (Helicoverpa armigera), Oriental leaf worm (Spodoptera litura), Pink bollworm (Pectinophora gossypiella), Cry1Ac resistant Pink bollworm (Pectinophora gossypiella), Soybean looper (Chrysodeixis includens), Southern armyworm (Spodoptera eridania), Southwestern corn borer (Diatraea grandiosella), Spotted bollworm (Earias vittella), Sugarcane borer (Diatraea saccharalis), Tobacco budworm (Heliothis virescens), and Velvet bean caterpillar (Anticarsia gemmatalis).

Also contemplated in this application are plants comprissegment encoding a pesticidal protein or fragment thereof, 35 ing a recombinant nucleic acid molecule comprising a heterologous promoter fragment operably linked to a polynucleotide segment encoding a pesticidal protein or fragment thereof, wherein: (a) said pesticidal protein comprises the amino acid sequence of SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (b) said pesticidal protein comprises an amino acid sequence having: (i) at least 93%, or 95%, or 98%, or 99%, or about 100% amino acid sequence identity to SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (ii) at least 73%, or 75%, or 80%, or 85%, or 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:8 or SEQ ID NO:10; or (c) said polynucleotide segment hybridizes under stringent hybridization conditions to the compliment of the nucleotide sequence of SEQ ID NO:3 or SEQ ID NO:9; or (d) said plant exhibits a detectable amount of said pesticidal protein. In certain embodiments, the pesticidal protein comprises SEQ ID NO:4, SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10. In one embodiment, the plant is either a dicotyledonous plant or a monocotyledonous plant. In another embodiment, the plant is further selected from the group consisting of an alfalfa, banana, barley, bean, broccoli, cabbage, brassica, carrot, cassava, castor, cauliflower, celery, chickpea, Chinese cabbage, citrus, coconut, coffee, corn, clover, cotton, a cucurbit, cucumber, Douglas fir, eggplant, eucalyptus, flax, garlic, grape, hops, leek, lettuce, Loblolly pine, millets, melons, nut, oat, olive, onion, ornamental, palm, pasture grass, pea, peanut, pepper, pigeon pea, pine, potato, poplar, pumpkin, Radiata pine, radish, rapeseed, rice, rootstocks, rye, safflower, shrub, sorghum, Southern pine, soybean, spinach, squash, strawberry, sugar beet, sugarcane, sunflower, sweet corn, sweet gum, sweet potato, switchgrass, tea, tobacco, tomato, triticale, turf grass, watermelon, and wheat.

In further embodiments, seeds comprising the recombinant nucleic acid molecules are disclosed.

In another embodiment, an insect inhibitory composition comprising the recombinant nucleic acid molecules disclosed in this application are contemplated. The insect inhibitory composition can further comprise a nucleotide sequence encoding at least one other pesticidal agent that is different from said pesticidal protein. In certain embodiments, the at least one other pesticidal agent is selected from the group consisting of an insect inhibitory protein, an insect inhibitory dsRNA molecule, and an ancillary protein. It is also contemplated that the at least one other pesticidal agent in the insect inhibitory composition exhibits activity against one or more pest species of the orders Lepidoptera, Coleoptera, or Hemiptera. The at least one other pesticidal agent in the insect inhibitory composition is in one embodiment selected from the group consisting of a Cry1A, Cry1Ab, Cry1Ac, Cry1A.105, Cry1Ae, Cry1B, Cry1C, Cry1C variants, Cry1D, Cry1E, Cry1F, Cry1A/F chimeras, Cry1G, 20 Cry1H, Cry1I, Cry1J, Cry1K, Cry1L, Cry2A, Cry2Ab, Cry2Ae, Cry3, Cry3A variants, Cry3B, Cry4B, Cry6, Cry7, Cry8, Cry9, Cry15, Cry34, Cry35, Cry43A, Cry43B, Cry51Aa1, ET29, ET33, ET34, ET35, ET66, ET70, TIC400, TIC407, TIC417, TIC431, TIC800, TIC807, TIC834, 25 TIC853, TIC900, TIC901, TIC1201, TIC1415, TIC2160, TIC3131, TIC836, TIC860, TIC867, TIC869, TIC1100, VIP3A, VIP3B, VIP3Ab, AXMI-AXMI-, AXMI-88, AXMI-97, AXMI-102, AXMI-112, AXMI-117, AXMI-100, AXMI-115, AXMI-113, and AXMI-005, AXMI134, AXMI-150, 30 AXMI-171, AXMI-184, AXMI-196, AXMI-204, AXMI-207, AXMI-209, AXMI-205, AXMI-218, AXMI-220, AXMI-221z, AXMI-222z, AXMI-223z, AXMI-224z and AXMI-225z, AXMI-238, AXMI-270, AXMI-279, AXMI-345, AXMI-335, AXMI-R1 and variants thereof, IP3 and 35 variants thereof, DIG-3, DIG-5, DIG-10, DIG-657 and a DIG-11 protein.

Commodity products comprising a detectable amount of the recombinant nucleic acid molecules disclosed in this application are also contemplated. Such commodity prod- 40 ucts include commodity corn bagged by a grain handler, corn flakes, corn cakes, corn flour, corn meal, corn syrup, corn oil, corn silage, corn starch, corn cereal, and the like, and corresponding soybean, rice, wheat, sorghum, pigeon pea, peanut, fruit, melon, and vegetable commodity products 45 including, where applicable, juices, concentrates, jams, jellies, marmalades, and other edible forms of such commodity products containing a detectable amount of such polynucleotides and or polypeptides of this application, whole or processed cotton seed, cotton oil, lint, seeds and plant parts 50 processed for feed or food, fiber, paper, biomasses, and fuel products such as fuel derived from cotton oil or pellets derived from cotton gin waste, whole or processed soybean seed, soybean oil, soybean protein, soybean meal, soybean flour, soybean flakes, soybean bran, soybean milk, soybean 55 cheese, soybean wine, animal feed comprising soybean, paper comprising soybean, cream comprising soybean, soybean biomass, and fuel products produced using soybean plants and soybean plant parts.

Also contemplated in this application are methods of 60 producing seed comprising the recombinant nucleic acid molecules disclosed in this application. The method comprises planting at least one of the seed comprising the recombinant nucleic acid molecules disclosed in this application; growing plant from the seed; and harvesting seed 65 from the plants, wherein the harvested seed comprises the recombinant nucleic acid molecules in this application.

6

In another illustrative embodiment, a plant resistant to insect infestation, is provided wherein the cells of said plant comprise: (a) a recombinant nucleic acid molecule encoding an insecticidally effective amount of a pesticidal protein as set forth in SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (b) an insecticidally effective amount of a protein comprising an amino acid sequence having: (i) at least 93%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (ii) at least 73%, or 75%, or 80%, or 85%, or 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:8 or SEQ ID NO:10.

Also disclosed in this application are methods for controlling a Lepidopteran species pest, and controlling a Lepidopteran species pest infestation of a plant, particularly a crop plant. The method comprises, in one embodiment, (a) contacting the pest with an insecticidally effective amount of a pesticidal proteins as set forth in SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (b) contacting the pest with an insecticidally effective amount of one or more pesticidal proteins comprising an amino acid sequence having: (i) at least 93%, or 95%, or about 100% amino acid sequence identity to identity to SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (ii) least 73%, or 75%, or 80%, or 85%, or 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:8 or SEQ ID NO:10.

Further provided herein is a method of detecting the presence of a recombinant nucleic acid molecule comprising a polynucleotide segment encoding a pesticidal protein or fragment thereof, wherein: (a) said pesticidal protein comprises the amino acid sequence of SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (b) said pesticidal protein comprises an amino acid sequence having: (i) at least 93%, or 95%, or 98%, or 99%, or about 100% amino acid sequence identity to SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (ii) at least 73%, or 75%, or 80%, or 85%, or 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:8 or SEQ ID NO:10; or (c) said polynucleotide segment hybridizes to a polynucleotide having the nucleotide sequence of SEQ ID NO:4, SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10. In one embodiment of the invention, the method comprises contacting a sample of nucleic acids with a nucleic acid probe that hybridizes under stringent hybridization conditions with genomic DNA from a plant comprising a polynucleotide segment encoding a pesticidal protein or fragment thereof provided herein, and does not hybridize under such hybridization conditions with genomic DNA from an otherwise isogenic plant that does not comprise the segment, wherein the probe is homologous or complementary to SEQ ID NO:3 or SEQ ID NO:9, or a sequence that encodes a pesticidal protein comprising an amino acid sequence having: (a) at least 93%, or 95%, or 98%, or 99%, or about 100% amino acid sequence identity to SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (b) at least 73%, or 75%, or 80%, or 85%, or 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:8 or SEQ ID NO:10. The method may further comprise (a) subjecting the sample and probe to stringent hybridization conditions; and (b) detecting hybridization of the probe with DNA of the sample.

Also provided by the invention are methods of detecting the presence of a pesticidal protein or fragment thereof in a sample comprising protein, wherein said pesticidal protein comprises the amino acid sequence of SEQ ID NO:4, SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, or SEQ ID NO:10; or said pesticidal protein comprises an amino acid sequence having: (a) at least 93%, or 95%, or 98%, or 99%, or about

100% amino acid sequence identity to SEQ ID NO:4, SEQ ID NO:2, or SEQ ID NO:6; or (b) at least 73%, or 75%, or 80%, or 85%, or 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:8 or SEQ ID NO:10. In one embodiment, the method comprises: (a) contacting a sample with an immunoreactive antibody; and (b) detecting the presence of the protein. In some embodiments the step of detecting comprises an ELISA, or a Western blot.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO:1 is a nucleic acid sequence encoding a TIC4472 pesticidal protein obtained from *Bacillus thuringiensis* species EG10742.

SEQ ID NO:2 is the amino acid sequence of the TIC4472 15 pesticidal protein.

SEQ ID NO:3 is a synthetic coding sequence encoding a TIC4472PL pesticidal protein designed for expression in a plant cell wherein an additional alanine codon is inserted immediately following the initiating methionine codon.

SEQ ID NO:4 is the amino acid sequence of TIC4472PL encoded by a synthetic coding sequence designed for expression in a plant cell (SEQ ID NO:3), and wherein an additional alanine amino acid is inserted immediately following the initiating methionine.

SEQ ID NO:5 is a nucleic acid sequence encoding a TIC1425 pesticidal protein obtained from *Bacillus thuringiensis* species EG10731.

SEQ ID NO:6 is the amino acid sequence of the TIC1425 pesticidal protein.

SEQ ID NO:7 is a nucleic acid sequence encoding a TIC2613 pesticidal protein obtained from *Bacillus thuringiensis* species EG5408.

SEQ ID NO:8 is the amino acid sequence of the TIC2613 pesticidal protein.

SEQ ID NO:9 is a synthetic coding sequence encoding a TIC2613PL pesticidal protein designed for expression in a plant cell wherein an additional alanine codon is inserted immediately following the initiating methionine codon.

SEQ ID NO:10 is the amino acid sequence of TIC2613PL 40 encoded by a synthetic coding sequence designed for expression in a plant cell (SEQ ID NO:9), and wherein an additional alanine amino acid is inserted immediately following the initiating methionine.

DETAILED DESCRIPTION OF THE INVENTION

The problem in the art of agricultural pest control can be characterized as a need for new toxin proteins that are 50 efficacious against target pests, exhibit broad spectrum toxicity against target pest species, are capable of being expressed in plants without causing undesirable agronomic issues, and provide an alternative mode of action compared to current toxins that are used commercially in plants.

Novel pesticidal proteins exemplified by TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL are disclosed herein, and address each of these needs, particularly against a broad spectrum of Lepidopteran insect pests, and more particularly against Beet armyworm (*Spodoptera* 60 exigua), Corn earworm (*Helicoverpa zea*), Cotton leaf worm (*Alabama argillacea*), European corn borer (*Ostrinia nubilalis*), Fall armyworm (*Spodoptera frugiperda*), Old World bollworm (*Helicoverpa armigera*), Oriental leaf worm (*Spodoptera litura*), Pink bollworm (*Pectinophora gossypiella*), 65 Cry1Ac resistant Pink bollworm (*Pectinophora gossypiella*), Soybean looper (*Chrysodeixis includens*), South-

8

ern armyworm (Spodoptera eridania), Southwestern corn borer (Diatraea grandiosella), Spotted bollworm (Earias vittella), Sugarcane borer (Diatraea saccharalis), Tobacco budworm (Heliothis virescens), and Velvet bean caterpillar (Anticarsia gemmatalis).

Reference in this application to TIC4472, "TIC4472 protein", "TIC4472 protein toxin", "TIC4472 toxin protein", "TIC4472 pesticidal protein", "TIC4472-related toxins", "TIC4472-related proteins", TIC4472PL, toxin 10 "TIC4472PL protein", "TIC4472PL protein toxin", "TIC4472PL toxin protein", "TIC4472PL pesticidal protein", "TIC4472PL-related toxins", "TIC4472PL-related toxin proteins", TIC1425, "TIC1425 protein", "TIC1425 protein toxin", "TIC1425 toxin protein", "TIC1425 pesticidal protein", "TIC1425 toxin protein", "TIC1425 related toxin proteins", "TIC2613, "TIC2613 protein", "TIC2613 protein", "TIC2613 protein toxin", "TIC2613 toxin protein", "TIC2613 related toxins", "TIC2613-related toxins" toxin proteins", TIC2613PL, "TIC2613PL protein", "TIC2613PL protein toxin", "TIC2613PL toxin protein", "TIC2613PL pesticidal protein", "TIC2613PL-related toxins", "TIC2613PL-related toxin proteins", and the like, refer to any novel pesticidal protein or insect inhibitory protein, that comprises, that consists of, that is substantially homologous to, that is similar to, or that is derived from any pesticidal protein or insect inhibitory protein sequence of TIC4472 (SEQ ID NO:2), TIC4472PL (SEQ ID NO:4), TIC1425 (SEQ ID NO:6), TIC2613 (SEQ ID NO:8), or TIC2613PL (SEQ ID NO:10) and pesticidal or insect inhibitory segments thereof, or combinations thereof, that confer activity against Lepidopteran pests, including any protein exhibiting pesticidal or insect inhibitory activity if alignment of such protein with TIC4472, TIC4472PL, or TIC1425 results in amino acid sequence identity of any fraction percentage from about 93% to about 100% percent; or if alignment of such protein with TIC2613 or TIC2613PL results in amino acid sequence identity of any fraction percentage from about 73% to about 100% percent. The TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL proteins include both the plastid-targeted and non-plastid targeted form of the proteins.

The term "segment" or "fragment" is used in this application to describe consecutive amino acid or nucleic acid sequences that are shorter than the complete amino acid or 45 nucleic acid sequence describing a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein. A segment or fragment exhibiting insect inhibitory activity is also disclosed in this application if alignment of such segment or fragment, with the corresponding section of the TIC4472 protein set forth in SEQ ID NO:2, TIC4472PL protein set forth in SEQ ID NO:4, TIC1425 protein set forth in SEQ ID NO:6, results in amino acid sequence identity of any fraction percentage from about 93 to about 100 percent between the segment or fragment and the corresponding section of the TIC4472, TIC4472PL, or TIC1425 protein; or if alignment of such segment or fragment, with the corresponding section of the TIC2613 set forth in SEQ ID NO:8, or TIC2613PL protein set forth in SEQ ID NO:10, results in amino acid sequence identity of any fraction percentage from about 73 to about 100 percent between the segment or fragment and the corresponding section of the TIC2613 or TIC2613PL protein.

In still further specific embodiments, a fragment of a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein may be defined as exhibiting pesticidal activity possessed by the starting protein molecule from which it is derived. A fragment of a nucleic acid sequence encoding a

TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein may be defined as encoding a protein exhibiting the pesticidal activity possessed by the protein molecule encoded by the starting nucleic acid sequence from which it is derived. A fragment or variant described herein may 5 further comprise a domain identified herein which is responsible for the pesticidal activity of a protein.

9

In specific embodiments, fragments of a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein are provided comprising at least about 50, at least about 75, at 10 least about 95, at least about 100, at least about 125, at least about 150, at least about 175, at least about 200, at least about 225, at least about 250, at least about 275, at least about 300, at least about 500, at least about 600, at least about 700, at least about 750, at least about 800, at least 15 about 900, at least about 1000, at least about 1100, at least about 1150, or at least about 1175 contiguous amino acids, or longer, of a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein having pesticidal activity as disclosed herein. In certain embodiments, the invention provides frag- 20 ments of any one of SEQ ID NOs: 2, 4, 6, 8, or 10, having the activity of the full length sequence. Methods for producing such fragments from a starting molecule are well known in the art.

Reference in this application to the terms "active" or 25 "activity", "pesticidal activity" or "pesticidal" or "insecticidal activity", "insect inhibitory" or "insecticidal" refer to efficacy of a toxic agent, such as a protein toxin, in inhibiting (inhibiting growth, feeding, fecundity, or viability), supity), controlling (controlling the pest infestation, controlling the pest feeding activities on a particular crop containing an effective amount of the TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein) or killing (causing the morbidity, mortality, or reduced fecundity of) a pest. These 35 terms are intended to include the result of providing a pesticidally effective amount of a toxic protein to a pest where the exposure of the pest to the toxic protein results in morbidity, mortality, reduced fecundity, or stunting. These terms also include repulsion of the pest from the plant, a 40 tissue of the plant, a plant part, seed, plant cells, or from the particular geographic location where the plant may be growing, as a result of providing a pesticidally effective amount of the toxic protein in or on the plant. In general, pesticidal activity refers to the ability of a toxic protein to be 45 effective in inhibiting the growth, development, viability, feeding behavior, mating behavior, fecundity, or any measurable decrease in the adverse effects caused by an insect feeding on this protein, protein fragment, protein segment or polynucleotide of a particular target pest, including but not 50 limited to insects of the order Lepidoptera. The toxic protein can be produced by the plant or can be applied to the plant or to the environment within the location where the plant is located. The terms "bioactivity", "effective", "efficacious" or variations thereof are also terms interchangeably utilized 55 in this application to describe the effects of proteins of the present invention on target insect pests.

A pesticidally effective amount of a toxic agent, when provided in the diet of a target pest, exhibits pesticidal activity when the toxic agent contacts the pest. A toxic agent 60 can be a pesticidal protein or one or more chemical agents known in the art. Pesticidal or insecticidal chemical agents and pesticidal or insecticidal protein agents can be used alone or in combinations with each other. Chemical agents include but are not limited to dsRNA molecules targeting 65 specific genes for suppression in a target pest, organochlorides, organophosphates, carbamates, pyrethroids, neonico10

tinoids, and ryanoids. Pesticidal or insecticidal protein agents include the protein toxins set forth in this application, as well as other proteinaceous toxic agents including those that target Lepidopterans, as well as protein toxins that are used to control other plant pests such as Cry and Cyt proteins available in the art for use in controlling Coleopteran, Hemipteran and Homopteran species.

It is intended that reference to a pest, particularly a pest of a crop plant, means insect pests of crop plants, particularly those Lepidoptera insect pests that are controlled by the TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein toxin class. However, reference to a pest can also include Coleopteran, Hemipteran and Homopteran insect pests of plants, as well as nematodes and fungi when toxic agents targeting these pests are co-localized or present together with the TIC4472, TIC4472PL, or TIC1425 protein or a protein that is 93 to about 100 percent identical to TIC4472, TIC4472PL, or TIC1425; or the TIC2613, or TIC2613PL protein or a protein that is 73 to about 100 percent identical to TIC2613, or TIC2613PL.

The TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL proteins are related by a common function and exhibit insecticidal activity towards insect pests from the Lepidoptera insect species, including adults, pupae, larvae, and neonates.

The insects of the order Lepidoptera include, but are not limited to, armyworms, cutworms, loopers, and heliothines in the Family Noctuidae, e.g., Fall armyworm (Spodoptera frugiperda), Beet armyworm (Spodoptera exigua), Black pressing (suppressing growth, feeding, fecundity, or viabil- 30 armyworm (Spodoptera exempta), Southern armyworm (Spodoptera eridania), bertha armyworm (Mamestra configurata), black cutworm (Agrotis ipsilon), cabbage looper (Trichoplusia ni), soybean looper (Pseudoplusia includens), velvetbean caterpillar (Anticarsia gemmatalis), green cloverworm (Hypena scabra), tobacco budworm (Heliothis virescens), granulate cutworm (Agrotis subterranea), armyworm (Pseudaletia unipuncta), western cutworm (Agrotis orthogonia); borers, casebearers, webworms, coneworms, cabbageworms and skeletonizers from the Family Pyralidae, e.g., European corn borer (Ostrinia nubilalis), navel orangeworm (Amyelois transitella), corn root webworm (Crambus caliginosellus), sod webworm (Herpetogramma licarsisalis), sunflower moth (Homoeosoma electellum), lesser cornstalk borer (Elasmopalpus lignosellus); leafrollers, budworms, seed worms, and fruit worms in the Family Tortricidae, e.g., codling moth (Cydia pomonella), grape berry moth (Endopiza viteana), oriental fruit moth (Grapholita molesta), sunflower bud moth (Suleima helianthana); and many other economically important Lepidoptera, e.g., diamondback moth (Plutella xylostella), pink bollworm (Pectinophora gossypiella), and gypsy moth (Lymantria dispar). Other insect pests of order Lepidoptera include, e.g., cotton leaf worm (Alabama argillacea), fruit tree leaf roller (Archips argyrospila), European leafroller (Archips rosana) and other Archips species, (Chilo suppressalis, Asiatic rice borer, or rice stem borer), rice leaf roller (Cnaphalocrocis medinalis), corn root webworm (Crambus caliginosellus), bluegrass webworm (Crambus teterrellus), southwestern corn borer (Diatraea grandiosella), surgarcane borer (Diatraea saccharalis), spiny bollworm (Earias insulana), spotted bollworm (Earias vittella), American bollworm (Helicoverpa armigera), corn earworm (Helicoverpa zea, also known as soybean podworm and cotton bollworm), tobacco budworm (Heliothis virescens), sod webworm (Herpetogramma licarsisalis), Western bean cutworm (Striacosta albicosta), European grape vine moth (Lobesia botrana), citrus leafminer (Phyllocnistis citrella),

large white butterfly (*Pieris brassicae*), small white butterfly (*Pieris rapae*, also known as imported cabbageworm), beet armyworm (*Spodoptera exigua*), tobacco cutworm (*Spodoptera litura*, also known as cluster caterpillar), and tomato leafminer (*Tuta absoluta*).

Reference in this application to an "isolated DNA molecule", or an equivalent term or phrase, is intended to mean that the DNA molecule is one that is present alone or in combination with other compositions, but not within its natural environment. For example, nucleic acid elements 10 such as a coding sequence, intron sequence, untranslated leader sequence, promoter sequence, transcriptional termination sequence, and the like, that are naturally found within the DNA of the genome of an organism are not considered to be "isolated" so long as the element is within the genome 15 of the organism and at the location within the genome in which it is naturally found. However, each of these elements, and subparts of these elements, would be "isolated" within the scope of this disclosure so long as the element is not within the genome of the organism and at the location within 20 the genome in which it is naturally found. Similarly, a nucleotide sequence encoding an insecticidal protein or any naturally occurring insecticidal variant of that protein would be an isolated nucleotide sequence so long as the nucleotide sequence was not within the DNA of the bacterium from 25 which the sequence encoding the protein is naturally found. A synthetic nucleotide sequence encoding the amino acid sequence of the naturally occurring insecticidal protein would be considered to be isolated for the purposes of this disclosure. For the purposes of this disclosure, any transgenic nucleotide sequence, i.e., the nucleotide sequence of the DNA inserted into the genome of the cells of a plant or bacterium, or present in an extrachromosomal vector, would be considered to be an isolated nucleotide sequence whether it is present within the plasmid or similar structure used to 35 transform the cells, within the genome of the plant or bacterium, or present in detectable amounts in tissues, progeny, biological samples or commodity products derived from the plant or bacterium.

As described further in this application, an open reading 40 frame (ORF) encoding TIC4747 (SEQ ID NO:1) was discovered in DNA obtained from Bacillus thuringiensis strain EG10742. The coding sequence was cloned and expressed in microbial host cells to produce recombinant proteins used in bioassays. An open reading frame (ORF) encoding TIC1425 (SEQ ID NO:5) was discovered in DNA obtained from Bacillus thuringiensis strain EG10731. An open reading frame (ORF) encoding TIC2613 (SEQ ID NO:7) was discovered in DNA obtained from Bacillus thuringiensis strain EG5408. Bioassay using microbial host cell-derived pro- 50 teins of TIC4472 demonstrated activity against the Lepidopteran species Beet armyworm (Spodoptera exigua), Corn earworm (Helicoverpa zea), Cotton leaf worm (Alabama argillacea), European corn borer (Ostrinia nubilalis), Fall armyworm (Spodoptera frugiperda), Old World bollworm 55 (Helicoverpa armigera), Oriental leaf worm (Spodoptera litura), Pink bollworm (Pectinophora gossypiella), Cry1Ac resistant Pink bollworm (Pectinophora gossypiella), Soybean looper (Chrysodeixis includens), Southern armyworm (Spodoptera eridania), Southwestern corn borer (Diatraea 60 grandiosella), Spotted bollworm (Earias vittella), Sugarcane borer (Diatraea saccharalis), Tobacco budworm (Heliothis virescens), and Velvet bean caterpillar (Anticarsia gemmatalis). In addition, activity was also observed against Yellow fever mosquito (Aedes aegypti). Bioassay using 65 microbial host cell-derived proteins of TIC1425 demonstrated activity against the Lepidopteran species Cotton leaf

12

worm (Alabama argillacea), European corn borer (Ostrinia nubilalis), Fall armyworm (Spodoptera frugiperda), Sugarcane borer (Diatraea saccharalis), and Southwestern corn borer (Diatraea grandiosella. Bioassay using microbial host cell-derived proteins of TIC2613 demonstrated activity against the Lepidopteran species Corn earworm (Helicoverpa zea), Cotton leaf worm (Alabama argillacea), European corn borer (Ostrinia nubilalis), Fall armyworm (Spodoptera frugiperda), Soybean looper (Chrysodeixis includens), Southwestern corn borer (Diatraea grandiosella), and Tobacco budworm (Heliothis virescens).

For expression in plant cells, the TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL proteins can be expressed to reside in the cytosol or targeted to various organelles of the plant cell. For example, targeting a protein to the chloroplast may result in increased levels of expressed protein in a transgenic plant while preventing off-phenotypes from occurring. Targeting may also result in an increase in pest resistance efficacy in the transgenic event. A target peptide or transit peptide is a short (3-70 amino acids long) peptide chain that directs the transport of a protein to a specific region in the cell, including the nucleus, mitochondria, endoplasmic reticulum (ER), chloroplast, apoplast, peroxisome and plasma membrane. Some target peptides are cleaved from the protein by signal peptidases after the proteins are transported. For targeting to the chloroplast, proteins contain transit peptides which are around 40-50 amino acids. For descriptions of the use of chloroplast transit peptides, see U.S. Pat. Nos. 5,188,642 and 5,728,925. Many chloroplast-localized proteins are expressed from nuclear genes as precursors and are targeted to the chloroplast by a chloroplast transit peptide (CTP). Examples of such isolated chloroplast proteins include, but are not limited to, those associated with the small subunit (SSU) of ribulose-1,5,bisphosphate carboxylase, ferredoxin, ferredoxin oxidoreductase, the light-harvesting complex protein I and protein II, thioredoxin F, enolpyruvyl shikimate phosphate synthase (EPSPS), and transit peptides described in U.S. Pat. No. 7,193,133. It has been demonstrated in vivo and in vitro that non-chloroplast proteins may be targeted to the chloroplast by use of protein fusions with a heterologous CTP and that the CTP is sufficient to target a protein to the chloroplast. Incorporation of a suitable chloroplast transit peptide such as the Arabidopsis thaliana EPSPS CTP (CTP2) (see, Klee et al., Mol. Gen. Genet. 210:437-442, 1987) or the Petunia hybrida EPSPS CTP (CTP4) (see, della-Cioppa et al., Proc. Natl. Acad. Sci. USA 83:6873-6877, 1986) has been shown to target heterologous EPSPS protein sequences to chloroplasts in transgenic plants (see, U.S. Pat. Nos. 5,627,061; 5,633,435; and 5,312,910; and EP 0218571; EP 189707; EP 508909; and EP 924299). For targeting the TIC6757 or TIC6757PL toxin protein to the chloroplast, a sequence encoding a chloroplast transit peptide is placed 5' in operable linkage and in frame to a synthetic coding sequence encoding the TIC6757 or TIC6757PL toxin protein that has been designed for optimal expression in plant cells.

It is contemplated that additional toxin protein sequences related to TIC4472, TIC1425, or TIC2613 can be created by using the amino acid sequence of TIC4472, TIC1425, or TIC2613 to create novel proteins with novel properties. The TIC4472, TIC1425, or TIC2613 toxin proteins can be aligned to combine differences at the amino acid sequence level into novel amino acid sequence variants and making appropriate changes to the recombinant nucleic acid sequence encoding the variants.

This disclosure further contemplates that improved variants of the TIC4472 protein toxin class can be engineered in planta by using various gene editing methods known in the art. Such technologies used for genome editing include, but are not limited to, ZFN (zinc-finger nuclease), meganucleases, TALEN (Transcription activator-like effector nucleases), and CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas (CRISPR-associated) systems. These genome editing methods can be used to alter the toxin protein coding sequence transformed within a plant cell to a different toxin coding sequence. Specifically, through these methods, one or more codons within the toxin coding sequence are altered to engineer a new protein amino acid sequence. Alternatively, a fragment within the coding sequence is replaced or deleted, or additional DNA fragments are inserted into the coding sequence, to engineer a new toxin coding sequence. The new coding sequence can encode a toxin protein with new properties such as increased activity or spectrum against insect pests, as well as provide activity against an insect pest species wherein resistance has developed against the original insect toxin protein. The plant cell comprising the gene edited toxin coding sequence can be used by methods known in the art to generate whole plants expressing the new toxin protein.

It is also contemplated that fragments of TIC4472, 25 TIC1425, or TIC2613 or protein variants thereof can be truncated forms wherein one or more amino acids are deleted from the N-terminal end, C-terminal end, the middle of the protein, or combinations thereof wherein the fragments and variants retain insect inhibitory activity. These fragments can be naturally occurring or synthetic variants of TIC4472, TIC1425, or TIC2613 or derived protein variants, but should retain the insect inhibitory activity of at least TIC4472, TIC1425, or TIC2613. A fragment or variant described herein may further comprise a domain identified herein which is responsible for the pesticidal activity of a protein.

14

are a result of a Clustal W alignment using these default parameters: Weight matrix: blosum, Gap opening penalty: 10.0, Gap extension penalty: 0.05, Hydrophilic gaps: On, Hydrophilic residues: GPSNDQERK, Residue-specific gap penalties: On (Thompson, et al (1994) Nucleic Acids Research, 22:4673-4680). Percent amino acid identity is further calculated by the product of 100% multiplied by (amino acid identities/length of subject protein). Other alignment algorithms are also available in the art and provide results similar to those obtained using a Clustal W alignment and are contemplated herein.

It is intended that a protein exhibiting insect inhibitory activity against a Lepidopteran insect species is related to TIC4472, TIC4472PL, or TIC1425 if the protein is used in a query, e.g., in a Clustal W alignment, and the proteins of the present invention as set forth as SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 are identified as hits in such alignment in which the query protein exhibits at least 93% to about 100% amino acid identity along the length of the query protein that is about 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, or any fraction percentage in this range. It is also intended that a protein exhibiting insect inhibitory activity against a Lepidopteran insect species is related to TIC2613 or TIC2613PL if the protein is used in a query, e.g., in a Clustal W alignment, and the proteins of the present invention as set forth as SEQ ID NO:8, or SEQ ID NO:10 are identified as hits in such alignment in which the query protein exhibits at least 73% to about 100% amino acid identity along the length of the query protein that is about 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, or any fraction percentage in this range

Exemplary proteins TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL were aligned with each other using a Clustal W algorithm. A pair-wise matrix of percent amino acid sequence identities for each of the full-length proteins was created, as reported in Table 1.

TABLE 1

	Pair-wise matrix display of exemplary proteins TIC4472, TIC4472PL, TIC1425, TIC2613. and TIC2613PL.									
Toxin	TIC4472 (SEQ ID NO: 2)	TIC4472PL (SEQ ID NO: 4)	TIC1425 (SEQ ID NO: 6)	TIC2613 (SEQ ID NO: 8)	TIC2613PL (SEQ ID NO: 10)					
TIC4472 (SEQ ID NO: 2)	_	99.9 (1186)	99.9 (1186)	68.1 (808)	68 (807)					
TIC4472PL (SEQ ID NO: 4)	99.8 (1186)	_	99.7 (1185)	67.9 (807)	68.1 (809)					
TIC1425 (SEQ ID NO: 6)	99.9 (1186)	99.8 (1185)	_	68.2 (809)	68.1 (808)					
TIC2613 (SEQ ID NO: 8)	68.6 (808)	68.5 (807)	68.7 (809)	_	99.9 (1177)					
TIC2613PL (SEQ ID NO: 10)	68.4 (807)	68.6 (809)	68.5 (808)	99.8 (1177)	_					

Table Description: Clustal W alignment between (X) and (Y) are reported in a pair-wise matrix. The percent amino acid identity between all pairs is calculated and is represented by the first number in each box. The second number (in parentheses) in each box represents the number of identical amino acids between the pair.

Proteins that resemble the TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL proteins can be identified and compared to each other using various computer 65 can also be related by primary structure (conserved amino based algorithms known in the art (see Tables 1 and 2). Amino acid sequence identities reported in this application

In addition to percent identity, TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL and related proteins acid motifs), by length (about 1187 amino acids), and by other characteristics. Characteristics of the TIC4472,

TIC4472PL, TIC1425, TIC2613, and TIC2613PL protein toxins are reported in Table 2.

16
plant leaf disks obtained from the transformed plants. Plants can be derived from the plant cells by regeneration, seed,

TABLE 2

	Selected characteris	stics of th	e TIC4472,	TIC4472F	L, TIC1425, T	IC2613, and T	IC2613PL prot	eins.
Protein	Molecular Weight (in Daltons)	Amino Acid Length	Isoelectric Point	Charge at PH 7.0	No. of Strongly Basic (-) Amino Acids	No. of Strongly Acidic Amino Acids	No. of Hydrophobic Amino Acids	No. of Polar Amino Acids
TIC4472 TIC4472F TIC1425 TIC2613 TIC2613F	134460.88 134636.07	1187 1188 1187 1178 1179	4.7545 4.7545 4.7545 4.6653 4.6653	-36 -36 -36 -42 -42	137 137 137 132 132	162 162 162 162 162	611 612 611 588 589	576 576 576 590 590

As described further in the Examples of this application, synthetic nucleic acid molecule sequences encoding a variant of TIC4472, TIC4472PL, and a variant of TIC2613, TIC2613PL, were designed for use in plants. An exemplary recombinant nucleic acid molecule sequence that was designed for use in plants encoding the TIC4472PL protein is presented as SEQ ID NO:3. An exemplary recombinant nucleic acid molecule sequence that was designed for use in $_{25}$ plants encoding the TIC2613PL protein is presented as SEQ ID NO:9. The TIC4472PL and TIC2613PL proteins have an additional alanine amino acid immediately following the initiating methionine relative to the TIC4472 and TIC2613 proteins, respectively. The additional alanine residue 30 inserted into the TIC4472 and TIC2613 amino acid sequences are believed to improve expression of the protein in planta. Likewise, synthetic nucleic acid molecule sequences encoding variants of TIC1425 and can designed for use in plants.

Leaf disc assay using $R_{\rm o}$ cotton leaf tissue expressing TIC4472PL protein demonstrated high activity against Soybean looper (*Chrysodeixis includens*) and Tobacco budworm (*Heliothis virescens*) and low activity against Cotton bollworm (*Helicoverpa zea*) and Fall armyworm (*Spodoptera frugiperda*). Leaf disc assay using $R_{\rm o}$ soybean leaf tissue expressing TIC4472PL protein demonstrated activity against Southern armyworm (*Spodoptera eridania*) and Soybean looper (*Chrysodeixis includens*).

Leaf samples from R₀ soybean plants expressing TIC4472PL and TIC2613PL proteins demonstrated activity against Southern armyworm (*Spodoptera eridania*) and Soybean looper (*Chrysodeixis includens*).

Expression cassettes and vectors containing a recombinant nucleic acid molecule sequence can be constructed and introduced into corn, soybean, cotton or other plant cells in accordance with transformation methods and techniques known in the art. For example, Agrobacterium-mediated transformation is described in U.S. Patent Application Pub- 55 lications 2009/0138985A1 (soybean), 2008/0280361A1 (soybean), 2009/0142837A1 (corn), 2008/0282432 (cotton), 2008/0256667 (cotton), 2003/0110531 (wheat), 2001/ 0042257 A1 (sugar beet), U.S. Pat. No. 5,750,871 (canola), U.S. Pat. No. 7,026,528 (wheat), and U.S. Pat. No. 6,365, 60 807 (rice), and in Arencibia et al. (1998) Transgenic Res. 7:213-222 (sugarcane) all of which are incorporated herein by reference in their entirety. Transformed cells can be regenerated into transformed plants that express TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL proteins 65 and demonstrate pesticidal activity through bioassays performed in the presence of Lepidopteran pest larvae using

pollen, or meristem transformation techniques. Methods for transforming plants are known in the art.

As an alternative to traditional transformation methods, a DNA sequence, such as a transgene, expression cassette(s), etc., may be inserted or integrated into a specific site or locus within the genome of a plant or plant cell via site-directed integration. Recombinant DNA construct(s) and molecule(s) of this disclosure may thus include a donor template sequence comprising at least one transgene, expression cassette, or other DNA sequence for insertion into the genome of the plant or plant cell. Such donor template for site-directed integration may further include one or two homology arms flanking an insertion sequence (i.e., the sequence, transgene, cassette, etc., to be inserted into the plant genome). The recombinant DNA construct(s) of this disclosure may further comprise an expression cassette(s) encoding a site-specific nuclease and/or any associated protein(s) to carry out site-directed integration. These nuclease expressing cassette(s) may be present in the same molecule or vector as the donor template (in cis) or on a separate molecule or vector (in trans). Several methods for site-directed integration are known in the art involving different proteins (or complexes of proteins and/or guide RNA) that cut the genomic DNA to produce a double strand break (DSB) or nick at a desired genomic site or locus. Briefly as understood in the art, during the process of repairing the DSB or nick introduced by the nuclease enzyme, the donor template DNA may become integrated into the genome at the site of the DSB or nick. The presence of the homology arm(s) in the donor template may promote the adoption and targeting of the insertion sequence into the plant genome during the repair process through homologous recombination, although an insertion event may occur through non-homologous end joining (NHEJ). Examples of site-specific nucleases that may be used include zinc-finger nucleases, engineered or native meganucleases, TALE-endonucleases, and RNA-guided endonucleases (e.g., Cas9 or Cpf1). For methods using RNA-guided site-specific nucleases (e.g., Cas9 or Cpf1), the recombinant DNA construct(s) will also comprise a sequence encoding one or more guide RNAs to direct the nuclease to the desired site within the plant genome.

As used herein, a "recombinant DNA molecule" is a DNA molecule comprising a combination of DNA molecules that would not naturally occur together without human intervention. For instance, a recombinant DNA molecule may be a DNA molecule that is comprised of at least two DNA molecules heterologous with respect to each other, a DNA molecule that comprises a DNA sequence that deviates from DNA sequences that exist in nature, or a DNA molecule that

has been incorporated into a host cell's DNA by genetic transformation or gene editing. Similarly, a "recombinant protein molecule" is a protein molecule comprising a combination of amino acids that would not naturally occur together without human intervention. For example, a recombinant protein molecule may be a protein molecule that is comprised of at least two amino acid molecules heterologous with respect to each other, a protein molecule that comprises an amino acid sequence that deviates from amino acid sequences that exist in nature, or a protein molecule that is expressed in a host cell as a result of genetic transformation of the host cell or by gene editing of the host cell genome.

Recombinant nucleic acid molecule compositions that encode TIC4472, TIC4472PL, TIC1425, TIC2613, and 15 TIC2613PL are contemplated. For example, TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL proteins can be expressed with recombinant DNA constructs in which a polynucleotide molecule with an ORF encoding the protein is operably linked to genetic expression elements 20 such as a promoter and any other regulatory element necessary for expression in the system for which the construct is intended. Non-limiting examples include a plant-functional promoter operably linked to a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein encoding 25 sequence for expression of the protein in plants or a Btfunctional promoter operably linked to a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein encoding sequence for expression of the protein in a Bt bacterium or other Bacillus species. Other elements can be 30 operably linked to the TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein encoding sequence including, but not limited to, enhancers, introns, untranslated leaders, encoded protein immobilization tags (HIS-tag), translocation peptides (i.e., plastid transit peptides, signal 35 peptides), polypeptide sequences for post-translational modifying enzymes, ribosomal binding sites, and RNAi target sites. Exemplary recombinant polynucleotide molecules provided herewith include, but are not limited to, a heterologous promoter operably linked to a polynucleotide 40 such as SEQ ID NO:3, SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:7, and SEQ ID NO:9 that encodes the respective polypeptides or proteins having the amino acid sequence as set forth in SEQ ID NO:4, SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, and SEQ ID NO:10. A heterologous pro- 45 moter can also be operably linked to synthetic DNA coding sequences encoding a plastid targeted TIC4472PL or TIC2613PL; or an untargeted TIC4472PL or TIC2613PL. The codons of a recombinant nucleic acid molecule encoding for proteins disclosed herein can be substituted by 50 synonymous codons (known in the art as a silent substitution).

A recombinant DNA construct comprising TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein encoding sequences can further comprise a region of DNA 55 that encodes for one or more insect inhibitory agents which can be configured to concomitantly express or co-express with a DNA sequence encoding a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein, a protein different from a TIC4472, TIC4472PL, TIC1425, TIC2613, or 60 TIC2613PL protein, an insect inhibitory dsRNA molecule, or an ancillary protein. Ancillary proteins include, but are not limited to, co-factors, enzymes, binding-partners, or other agents that function to aid in the effectiveness of an insect inhibitory agent, for example, by aiding its expression, influencing its stability in plants, optimizing free energy for oligomerization, augmenting its toxicity, and

18

increasing its spectrum of activity. An ancillary protein may facilitate the uptake of one or more insect inhibitory agents, for example, or potentiate the toxic effects of the toxic agent.

A recombinant DNA construct can be assembled so that all proteins or dsRNA molecules are expressed from one promoter or each protein or dsRNA molecules is under separate promoter control or some combination thereof. The proteins of this invention can be expressed from a multigene expression system in which one or more proteins of TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL are expressed from a common nucleotide segment which also contains other open reading frames and promoters, depending on the type of expression system selected. For example, a bacterial multi-gene expression system can utilize a single promoter to drive expression of multiply-linked/ tandem open reading frames from within a single operon (i.e., polycistronic expression). In another example, a plant multi-gene expression system can utilize multiply-unlinked or linked expression cassettes, each cassette expressing a different protein or other agent such as one or more dsRNA molecules.

Recombinant polynucleotides or recombinant DNA constructs comprising a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein encoding sequence can be delivered to host cells by vectors, e.g., a plasmid, baculovirus, synthetic chromosome, virion, cosmid, phagemid, phage, or viral vector. Such vectors can be used to achieve stable or transient expression of a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein encoding sequence in a host cell, or subsequent expression of the encoded polypeptide. An exogenous recombinant polynucleotide or recombinant DNA construct that comprises a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein encoding sequence and that is introduced into a host cell is referred in this application as a "transgene".

Transgenic bacteria, transgenic plant cells, transgenic plants, and transgenic plant parts that contain a recombinant polynucleotide that expresses any one or more of TIC4472 or a related family toxin protein encoding sequence are provided herein. The term "bacterial cell" or "bacterium" can include, but is not limited to, an Agrobacterium, a Bacillus, an Escherichia, a Salmonella, a Pseudomonas, a Brevibacillus, a Klebsiella, an Erwinia, or a Rhizobium cell. The term "plant cell" or "plant" can include but is not limited to a dicotyledonous or monocotyledonous plant. The term "plant cell" or "plant" can also include but is not limited to an alfalfa, banana, barley, bean, broccoli, cabbage, brassica, carrot, cassava, castor, cauliflower, celery, chickpea, Chinese cabbage, citrus, coconut, coffee, corn, clover, cotton, a cucurbit, cucumber, Douglas fir, eggplant, eucalyptus, flax, garlic, grape, hops, leek, lettuce, Loblolly pine, millets, melons, nut, oat, olive, onion, ornamental, palm, pasture grass, pea, peanut, pepper, pigeonpea, pine, potato, poplar, pumpkin, Radiata pine, radish, rapeseed, rice, rootstocks, rye, safflower, shrub, sorghum, Southern pine, soybean, spinach, squash, strawberry, sugar beet, sugarcane, sunflower, sweet corn, sweet gum, sweet potato, switchgrass, tea, tobacco, tomato, triticale, turf grass, watermelon, and wheat plant cell or plant. In certain embodiments, transgenic plants and transgenic plant parts regenerated from a transgenic plant cell are provided. In certain embodiments, the transgenic plants can be obtained from a transgenic seed, by cutting, snapping, grinding or otherwise disassociating the part from the plant. In certain embodiments, the plant part can be a seed, a boll, a leaf, a flower, a stem, a root, or any portion thereof, or a non-regenerable portion of a transgenic plant part. As used in this context, a "non-

regenerable" portion of a transgenic plant part is a portion that cannot be induced to form a whole plant or that cannot be induced to form a whole plant that is capable of sexual and/or asexual reproduction. In certain embodiments, a non-regenerable portion of a plant part is a portion of a 5 transgenic seed, boll, leaf, flower, stem, or root.

Methods of making transgenic plants that comprise insect, Lepidoptera-inhibitory amounts of a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein are provided. Such plants can be made by introducing a recombinant 10 polynucleotide that encodes any of the proteins provided in this application into a plant cell, and selecting a plant derived from said plant cell that expresses an insect, Lepidoptera-inhibitory amount of the proteins. Plants can be derived from the plant cells by regeneration, seed, pollen, or 15 meristem transformation techniques. Methods for transforming plants are known in the art.

Processed plant products, wherein the processed product comprises a detectable amount of a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein, an insect inhibi- 20 tory segment or fragment thereof, or any distinguishing portion thereof, are also disclosed herein. In certain embodiments, the processed product is selected from the group consisting of plant parts, plant biomass, oil, meal, sugar, animal feed, flour, flakes, bran, lint, hulls, processed seed, 25 and seed. In certain embodiments, the processed product is non-regenerable. The plant product can comprise commodity or other products of commerce derived from a transgenic plant or transgenic plant part, where the commodity or other products can be tracked through commerce by detecting 30 nucleotide segments or expressed RNA or proteins that encode or comprise distinguishing portions of a TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein.

Plants expressing the TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL proteins can be crossed by breed- 35 ing with transgenic events expressing other toxin proteins and/or expressing other transgenic traits such as herbicide tolerance genes, genes conferring yield or stress tolerance traits, and the like, or such traits can be combined in a single vector so that the traits are all linked.

Reference in this application to an "isolated" DNA molecule or amino acid molecule, or an equivalent term or phrase, is intended to mean that the DNA molecule or amino acid molecule is one that is present alone or in combination with other compositions, but not within its natural environ- 45 ment. For example, a DNA molecule or amino acid molecule would be "isolated" within the scope of this disclosure so long as the element is not within the genome of the organism and at the location within the genome in which it is naturally found. For the purposes of this disclosure, any transgenic 50 nucleotide sequence, i.e., the nucleotide sequence of the DNA inserted into the genome of the cells of a plant or bacterium, or present in an extrachromosomal vector, would be considered to be an isolated nucleotide sequence whether it is present within the plasmid or similar structure used to 55 transform the cells, within the genome of the plant or bacterium, or present in detectable amounts in tissues, progeny, biological samples or commodity products derived from the plant or bacterium.

As further described in the Examples, TIC4472, 60 TIC4472PL, TIC1425, TIC2613, or TIC2613PL protein-encoding sequences and sequences having a substantial percentage identity to TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL can be identified using methods known to those of ordinary skill in the art such as polymerase chain reaction (PCR), thermal amplification and hybridization. For example, the proteins TIC4472,

20

TIC4472PL, TIC1425, TIC2613, or TIC2613PL can be used to produce antibodies that bind specifically to related proteins, and can be used to screen for and to find other protein members that are closely related.

Furthermore, nucleotide sequences encoding the TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL toxin proteins can be used as probes and primers for screening to identify other members of the class using thermalcycle or isothermal amplification and hybridization methods. For example, oligonucleotides derived from the sequence as set forth in SEQ ID NO:3 can be used to determine the presence or absence of a TIC4472PL transgene in a deoxyribonucleic acid sample derived from a commodity product. Oligonucleotides derived from the sequence as set forth in SEQ ID NO:7 can be used to determine the presence or absence of a TIC2613PL transgene in a deoxyribonucleic acid sample derived from a commodity product. Given the sensitivity of certain nucleic acid detection methods that employ oligonucleotides, it is anticipated that oligonucleotides derived from sequences as set forth in SEQ ID NO:3 or SEQ ID NO:9 can be used to detect a TIC4472PL or TIC2613PL transgene in commodity products derived from pooled sources where only a fraction of the commodity product is derived from a transgenic plant containing any of the transgenes. It is further recognized that such oligonucleotides can be used to introduce nucleotide sequence variation in each of SEQ ID NO:3 and SEQ ID NO:9. Such "mutagenesis" oligonucleotides are useful for identification of TIC4472PL and TIC2613PL amino acid sequence variants exhibiting a range of insect inhibitory activity or varied expression in transgenic plant host cells.

Nucleotide sequence homologs, e.g., insecticidal proteins encoded by nucleotide sequences that hybridize to each or any of the sequences disclosed in this application under stringent hybridization conditions, are also an embodiment of the present invention. The invention also provides a method for detecting a first nucleotide sequence that hybridizes to a second nucleotide sequence, wherein the first nucleotide sequence (or its reverse complement sequence) 40 encodes a pesticidal protein or pesticidal fragment thereof and hybridizes to the second nucleotide sequence. In such case, the second nucleotide sequence can be any of the nucleotide sequences presented as SEQ ID NO:3, SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:9 under stringent hybridization conditions. Nucleotide coding sequences hybridize to one another under appropriate hybridization conditions, such as stringent hybridization conditions, and the proteins encoded by these nucleotide sequences cross react with antiserum raised against any one of the other proteins. Stringent hybridization conditions, as defined herein, comprise hybridization at a temperature of at least 42° C. followed by two washes for five minutes each at room temperature with 2×SSC, 0.1% SDS, followed by two washes for thirty minutes each at 65° C. in 0.5×SSC, 0.1% SDS. Washes at even higher temperatures constitute even more stringent conditions, e.g., hybridization conditions of 68° C., followed by washing at 68° C., in 2×SSC containing 0.1% SDS.

One skilled in the art will recognize that, due to the redundancy of the genetic code, many other sequences are capable of encoding such related proteins, and those sequences, to the extent that they function to express pesticidal proteins either in *Bacillus* strains or in plant cells, are embodiments of the present invention, recognizing of course that many such redundant coding sequences will not hybridize under these conditions to the native *Bacillus* sequences encoding TIC4472, TIC4472PL, TIC1425, TIC2613, and

TIC2613PL. This application contemplates the use of these and other identification methods known to those of ordinary skill in the art, to identify TIC4472, TIC1425, and TIC2613 protein-encoding sequences and sequences having a substantial percentage identity to TIC4472, TIC1425, and 5 TIC2613 protein-encoding sequences.

This disclosure also contemplates the use of molecular methods known in the art to engineer and clone commercially useful proteins comprising chimeras of proteins from pesticidal proteins; e.g., the chimeras may be assembled from segments of the TIC4472, TIC4472PL, TIC1425, TIC2613, or TIC2613PL, proteins to derive additional useful embodiments including assembly of segments of TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL proteins with segments of diverse proteins different from 15 TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL and related proteins. The TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL proteins may be subjected to alignment to each other and to other Bacillus thuringiensis or other pesticidal proteins (whether or not these are closely 20 or distantly related phylogenetically), and segments of each such protein may be identified that are useful for substitution between the aligned proteins, resulting in the construction of chimeric proteins. Such chimeric proteins can be subjected to pest bioassay analysis and characterized for the presence 25 or absence of increased bioactivity or expanded target pest spectrum compared to the parent proteins from which each such segment in the chimera was derived. The pesticidal activity of the polypeptides may be further engineered for activity to a particular pest or to a broader spectrum of pests 30 by swapping domains or segments with other proteins or by using directed evolution methods known in the art.

Methods of controlling insects, in particular Lepidoptera infestations of crop plants, with the TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL proteins are also dis- 35 closed in this application. Such methods can comprise growing a plant comprising an insect- or Lepidopterainhibitory amount of a TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL toxin protein. In certain embodiments, such methods can further comprise any one or more 40 of: (i) applying any composition comprising or encoding a TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL toxin protein to a plant or a seed that gives rise to a plant; and (ii) transforming a plant or a plant cell that gives rise to a plant with a polynucleotide encoding a TIC4472, 45 TIC4472PL, TIC1425, TIC2613, and TIC2613PL toxin protein. In general, it is contemplated that a TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL toxin protein can be provided in a composition, provided in a microorganism, or provided in a transgenic plant to confer insect 50 inhibitory activity against Lepidopteran insects.

In certain embodiments, a recombinant nucleic acid molecule of TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL toxin proteins is the insecticidally active ingredient of an insect inhibitory composition prepared by cul- 55 turing recombinant Bacillus or any other recombinant bacterial cell transformed to express a TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL toxin protein under conditions suitable to express the TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL toxin protein. Such a 60 composition can be prepared by desiccation, lyophilization, homogenization, extraction, filtration, centrifugation, sedimentation, or concentration of a culture of such recombinant cells expressing/producing said recombinant polypeptide. Such a process can result in a Bacillus or other entomopatho- 65 genic bacterial cell extract, cell suspension, cell homogenate, cell lysate, cell supernatant, cell filtrate, or cell pellet.

22

By obtaining the recombinant polypeptides so produced, a composition that includes the recombinant polypeptides can include bacterial cells, bacterial spores, and parasporal inclusion bodies and can be formulated for various uses, including as agricultural insect inhibitory spray products or as insect inhibitory formulations in diet bioassays.

In one embodiment, to reduce the likelihood of resistance development, an insect inhibitory composition comprising TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL can further comprise at least one additional polypeptide that exhibits insect inhibitory activity against the same Lepidopteran insect species, but which is different from the TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL toxin protein. Possible additional polypeptides for such a composition include an insect inhibitory protein and an insect inhibitory dsRNA molecule. One example for the use of such ribonucleotide sequences to control insect pests is described in Baum, et al. (U.S. Patent Publication 2006/ 0021087 A1). Such additional polypeptide for the control of Lepidopteran pests may be selected from the group consisting of an insect inhibitory protein, such as, but not limited to, Cry1A (U.S. Pat. No. 5,880,275), Cry1Ab, Cry1Ac, Cry1A.105, Cry1Ae, Cry1B (U.S. patent Ser. No. 10/525, 318), Cry1C (U.S. Pat. No. 6,033,874), Cry1D, Cry1Da and variants thereof, Cry1E, Cry1F, and Cry1A/F chimeras (U.S. Pat. Nos. 7,070,982; 6,962,705; and 6,713,063), Cry1G, Cry1H, Cry1l, Cry1J, Cry1K, Cry1L, Cry1-type chimeras such as, but not limited to, TIC836, TIC860, TIC867, TIC869, and TIC1100 (International Application Publication WO2016/061391 (A2)), TIC2160 (International Application Publication WO2016/061392(A2)), Cry2A, Cry2Ab (U.S. Pat. No. 7,064,249), Cry2Ae, Cry4B, Cry6, Cry7, Cry8, Cry9, Cry15, Cry43A, Cry43B, Cry51Aa1, ET66, TIC400, TIC800, TIC834, TIC1415, Vip3A, VIP3Ab, VIP3B, AXMI-001, AXMI-002, AXMI-030, AXMI-035, AND AXMI-045 (U.S. Patent Publication 2013-0117884 A1), AXMI-52, AXMI-58, AXMI-88, AXMI-97, AXMI-102, AXMI-112, AXMI-117, AXMI-100 (U.S. Patent Publication 2013-0310543 A1), AXMI-115, AXMI-113, AXMI-005 (U.S. Patent Publication 2013-0104259 A1), AXMI-134 (U.S. Patent Publication 2013-0167264 A1), AXMI-150 (U.S. Patent Publication 2010-0160231 A1), AXMI-184 (U.S. Patent Publication 2010-0004176 A1), AXMI-196, AXMI-204, AXMI-207, axmi209 (U.S. Patent Publication 2011-0030096 A1), AXMI-218, AXMI-220 (U.S. Patent Publication 2014-0245491 A1), AXMI-221z, AXMI-222z, AXMI-223z, AXMI-224z, AXMI-225z (U.S. Patent Publication 2014-0196175 A1), AXMI-238 (U.S. Patent Publication 2014-0033363 A1), AXMI-270 (U.S. Patent Publication 2014-0223598 A1), AXMI-345 (U.S. Patent Publication 2014-0373195 A1), AXMI-335 (International Application Publication WO2013/134523(A2)), DIG-3 (U.S. Patent Publication 2013-0219570 A1), DIG-5 (U.S. Patent Publication 2010-0317569 A1), DIG-11 (U.S. Patent Publication 2010-0319093 A1), AfIP-1A and derivatives thereof (U.S. Patent Publication 2014-0033361 A1), AfIP-1B and derivatives thereof (U.S. Patent Publication 2014-0033361 A1), PIP-1APIP-1B (U.S. Patent Publication 2014-0007292 A1), PSEEN3174 (U.S. Patent Publication 2014-0007292 A1), AECFG-592740 (U.S. Patent Publication 2014-0007292 A1), Pput_1063 (U.S. Patent Publication 2014-0007292 A1), DIG-657 (International Application Publication WO2015/195594(A2)), Pput_1064 (U.S. Patent Publication 2014-0007292 A1), GS-135 and derivatives thereof (U.S. Patent Publication 2012-0233726 A1), GS153 and derivatives thereof (U.S. Patent Publication 2012-0192310 A1), GS154 and derivatives thereof (U.S. Patent Publication

2012-0192310 A1), GS155 and derivatives thereof (U.S. Patent Publication 2012-0192310 A1), SEQ ID NO:2 and derivatives thereof as described in U.S. Patent Publication 2012-0167259 A1, SEQ ID NO:2 and derivatives thereof as described in U.S. Patent Publication 2012-0047606 A1, SEQ 5 ID NO:2 and derivatives thereof as described in U.S. Patent Publication 2011-0154536 A1, SEQ ID NO:2 and derivatives thereof as described in U.S. Patent Publication 2011-0112013 A1, SEQ ID NO:2 and 4 and derivatives thereof as described in U.S. Patent Publication 2010-0192256 A1, SEQ 10 ID NO:2 and derivatives thereof as described in U.S. Patent Publication 2010-0077507 A1, SEQ ID NO:2 and derivatives thereof as described in U.S. Patent Publication 2010-0077508 A1, SEQ ID NO:2 and derivatives thereof as described in U.S. Patent Publication 2009-0313721 A1, SEQ ID NO:2 or 4 and derivatives thereof as described in U.S. Patent Publication 2010-0269221 A1, SEQ ID NO:2 and derivatives thereof as described in U.S. Pat. No. 7,772,465 (B2), CF161 0085 and derivatives thereof as described in WO2014/008054 A2, Lepidopteran toxic proteins and their 20 derivatives as described in US Patent Publications US2008-0172762 A1, US2011-0055968 A1, and US2012-0117690 A1; SEQ ID NO:2 and derivatives thereof as described in U.S. Pat. No. 7,510,878(B2), SEQ ID NO:2 and derivatives thereof as described in U.S. Pat. No. 7,812,129(B1); and the 25

In other embodiments, such composition/formulation can further comprise at least one additional polypeptide that exhibits insect inhibitory activity to an insect that is not inhibited by an otherwise insect inhibitory protein of the 30 present invention to expand the spectrum of insect inhibition obtained. For example, for the control of Hemipteran pests, combinations of insect inhibitory proteins of the present invention can be used with Hemipteran-active proteins such as TIC1415 (US Patent Publication 2013-0097735 A1), 35 TIC807 (U.S. Pat. No. 8,609,936), TIC834 (U.S. Patent Publication 2013-0269060 A1), AXMI-036 (U.S. Patent Publication 2010-0137216 A1), and AXMI-171 (U.S. Patent Publication 2013-0055469 A1). Further a polypeptide for the control of Coleopteran pests may be selected from the 40 group consisting of an insect inhibitory protein, such as, but not limited to, Cry3Bb (U.S. Pat. No. 6,501,009), Cry1C variants, Cry3A variants, Cry3, Cry3B, Cry34/35, 5307, AXMI134 (U.S. Patent Publication 2013-0167264 A1) AXMI-184 (U.S. Patent Publication 2010-0004176 A1), 45 AXMI-205 (U.S. Patent Publication 2014-0298538 A1), axmi207 (U.S. Patent Publication 2013-0303440 A1). AXMI-218, AXMI-220 (U.S. Patent 20140245491A1), AXMI-221z, AXMI-223z (U.S. Patent Publication 2014-0196175 A1), AXMI-279 (U.S. Patent 50 Publication 2014-0223599 A1), AXMI-R1 and variants thereof (U.S. Patent Publication 2010-0197592 A1, TIC407, TIC417, TIC431, TIC807, TIC853, TIC901, TIC1201, TIC3131, DIG-10 (U.S. Patent Publication 2010-0319092 A1), eHIPs (U.S. Patent Application Publication No. 2010/ 55 0017914). IP3 and variants thereof (U.S. Patent Publication 2012-0210462 A1), and $\overline{\omega}$ -Hexatoxin-Hv1a (U.S. Patent Application Publication US2014-0366227 A1).

Additional polypeptides for the control of *Coleopteran*, Lepidopteran, and *Hemipteran* insect pests can be found on 60 the *Bacillus thuringiensis* toxin nomenclature website maintained by Neil Crickmore (on the world wide web at btnomenclature.info).

The possibility for insects to develop resistance to certain insecticides has been documented in the art. One insect 65 resistance management strategy is to employ transgenic crops that express two distinct insect inhibitory agents that

24

operate through different modes of action. Therefore, any insects with resistance to either one of the insect inhibitory agents can be controlled by the other insect inhibitory agent. Another insect resistance management strategy employs the use of plants that are not protected to the targeted Lepidopteran pest species to provide a refuge for such unprotected plants. One particular example is described in U.S. Pat. No. 6,551,962, which is incorporated by reference in its entirety.

Other embodiments such as topically applied pesticidal chemistries that are designed for controlling pests that are also controlled by the proteins disclosed herein to be used with proteins in seed treatments, spray on, drip on, or wipe on formulations can be applied directly to the soil (a soil drench), applied to growing plants expressing the proteins disclosed herein, or formulated to be applied to seed containing one or more transgenes encoding one or more of the proteins disclosed. Such formulations for use in seed treatments can be applied with various stickers and tackifiers known in the art. Such formulations can contain pesticides that are synergistic in mode of action with the proteins disclosed, so that the formulation pesticides act through a different mode of action to control the same or similar pests that can be controlled by the proteins disclosed, or that such pesticides act to control pests within a broader host range or plant pest species that are not effectively controlled by the TIC4472, TIC4472PL, TIC1425, TIC2613, and TIC2613PL pesticidal proteins.

The aforementioned composition/formulation can further comprise an agriculturally-acceptable carrier, such as a bait, a powder, dust, pellet, granule, spray, emulsion, a colloidal suspension, an aqueous solution, a *Bacillus* spore/crystal preparation, a seed treatment, a recombinant plant cell/plant tissue/seed/plant transformed to express one or more of the proteins, or bacterium transformed to express one or more of the proteins. Depending on the level of insect inhibitory or insecticidal inhibition inherent in the recombinant polypeptide and the level of formulation to be applied to a plant or diet assay, the composition/formulation can include various by weight amounts of the recombinant polypeptide, e.g. from 0.0001% to 0.001% to 0.01% to 1% to 99% by weight of the recombinant polypeptide.

In view of the foregoing, those of skill in the art should appreciate that changes can be made in the specific aspects which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention. Thus, specific structural and functional details disclosed herein are not to be interpreted as limiting. It should be understood that the entire disclosure of each reference cited herein is incorporated within the disclosure of this application.

EXAMPLES

Example 1

Discovery, Cloning, and Expression of TIC4472, TIC1425, and TIC2613

Sequences encoding three novel *Bacillus thuringiensis* pesticidal proteins were identified, cloned, sequence confirmed, and tested in insect bioassay. The pesticidal proteins, TIC4472, TIC1425 and TIC2613, isolated from the *Bacillus thuringiensis* strains EG10742, EG10731, and EG5408, respectively, represent novel Cry1Ca-like proteins.

Polymerase chain reaction (PCR) primers were designed to amplify a full length copy of the coding region for

26

TIC4472, TIC1425 and TIC2613 from total genomic DNA isolated from the *Bacillus thuringiensis* strains EG10742, EG10731, and EG5408, respectively. The PCR amplicons also included the translational initiation and termination codons of each coding sequence.

Each of the amplicons were cloned using methods known in the art into Bt (*Bacillus thuringiensis*) expression vectors in operable linkage with a Bt expressible promoter.

Example 2

TIC4472, TIC1425, and TIC2613 Demonstrates Lepidopteran Activity in Insect Bioassay

Bioactivity of the pesticidal proteins TIC4472, TIC1425 and TIC2613 was evaluated by producing the protein in a Bt expression host. A Bt strain expressing TIC4472, TIC1425 and TIC2613 was grown for twenty four (24) hours and then either a spore crystal preparation or solubilized protein preparation was added to insect diet. Mortality and stunting were evaluated by comparing the growth and development of insects on a diet with a culture from the Bt strain expressing TIC4472, TIC1425 and TIC2613 to insects on a diet with an untreated control culture.

Preparations of TIC4472 were assayed against the Lepidopteran species Corn earworm (two colonies (CEW and CEWUC), Helicoverpa zea, also herein referred to as Cotton bollworm and Soybean pod worm), Cotton leaf worm (CLW, Alabama argillacea), European corn borer (ECB, Ostrinia nubilalis), Fall armyworm (FAW, Spodoptera frugiperda), 30 Soybean looper (SBL, Chrysodeixis includens), Southern armyworm (SAW, Spodoptera eridania), Southwestern corn borer (SWCB, Diatraea grandiosella), Sugarcane borer (SCB, Diatraea saccharalis), a Cry2Ab resistant colony of Sugarcane borer (SCB2R), Tobacco budworm (TBW, 35 Heliothis virescens), and Velvet bean caterpillar (VBW, Anticarsia gemmatalis); the Coleopteran species Colorado potato beetle (CPB, Leptinotarsa decemlineata) and West-

Agrotis ipsilon), Cotton leaf worm (CLW, Alabama argillacea), European corn borer (ECB, Ostrinia nubilalis), Fall armyworm (FAW, Spodoptera frugiperda), Black cutworm (BCW, Agrotis ipsilon), Southwestern corn borer (SWCB, Diatraea grandiosella), and Sugarcane borer (SCB, Diatraea saccharalis); and the Coleopteran species Western corn rootworm (WCB, Diabrotica virgifera virgifera) and Southern corn rootworm (Diabrotica undecimpunctata howardi). Preparations of TIC2613 were assayed against the Lepidopteran species Corn earworm (two colonies (CEW and CEWUC), Helicoverpa zea, also herein referred to as Cotton bollworm and Soybean pod worm), Cotton leaf worm (CLW, Alabama argillacea), European corn borer (ECB, Ostrinia nubilalis), Fall armyworm (FAW, Spodoptera frugiperda), Soybean looper (SBL, Chrysodeixis includens), Southern armyworm (SAW, Spodoptera eridania), Black cutworm (BCW, Agrotis ipsilon), Southwestern corn borer (SWCB, Diatraea grandiosella), Tobacco budworm (TBW, Heliothis virescens), and Velvet bean caterpillar (VBW, Anticarsia gemmatalis); the Coleopteran species Colorado potato beetle (CPB, Leptinotarsa decemlineata); Western corn rootworm (WCB, Diabrotica virgifera virgifera) and Southern corn rootworm (Diabrotica undecimpunctata howardi); and the Hemipteran species Tarnished plant bug (TPB, Lygus lineolaris), Western tarnished plant bug (WTP, Lygus hesperus).

The bioassay activity observed for each protein grown in the Bt host is presented in Tables 3 and 4 below, wherein "+" indicates activity, "NT" indicates the toxin was not assayed against that specific insect pest, "S" indicates stunting, and "M" indicates mortality. Preparations of TIC4472, TIC1425, and TIC2613 did not demonstrate activity against the *Coleopteran* or the *Hemipteran* insect pests assayed for each protein. TIC4472 also demonstrated activity against Yellow fever mosquito (*Aedes aegypti*). All three toxins demonstrated resistance to multiple Lepidopteran insect pests as shown in Tables 3 and 4.

TABLE 3

Bioas	ssay act	tivity o	f TIC4	472, TI	C142	5, an	d TI	C261:	3 aga	ainst	insec	t pest	s.	
	C	EW_	CEV	<u>vuc</u>	CI	<u>.w</u>	E	CB	FA	<u>w</u>	SI	<u>3L</u> .	SA	W
Toxin	S	M	S	M	S	M	S	M	s	M	S	M	s	M
TIC4472 TIC1425 TIC2613	+ NT +	+ NT	+	+	+	+++++	+++++	+++++	+++++	+	+ NT +	+ NT +	+	+

TABLE 4

Bios	assay act	ivity c	of TIC	4472,	TIC14	125, ar	nd TI	C261	3 aga	inst	insec	t pes	ts.	
	B0	CW_	S	СВ	SC	B2R	SW	CB	TE	8W	VI	<u>3C</u>	YI	FM_
Toxin	S	M	S	M	S	M	S	M	S	M	S	M	S	M
TIC4472 TIC1425 TIC2613	NT	NT	+	+ + +	+ NT NT	+ NT NT	+	++++	+ NT +	+ NT	+ NT +	+ NT	NT NT	+ NT NT

ern corn rootworm (WCB, Diabrotica virgifera virgifera); the Hemipteran species Tarnished plant bug (TPB, Lygus lineolaris), Western tarnished plant bug (WTP, Lygus hesperus); and the Dipteran species Yellow fever mosquito 65 (Aedes aegypti). Preparations of TIC1425 were assayed against the Lepidopteran species Black cutworm (BCW,

As can be seen in Tables 3 and 4 above, the insect toxin TIC4472 demonstrated activity against all Lepidopteran insect pests assayed (CEW, CEWUC, CLW, ECB, FAW, SBL, SAW, SCB, SCB2R, SWCB, TBW, and VBC); as well as YFM. The insect toxin TIC1425 demonstrated activity against CLW, ECB, FAW, SCB, and SWCB. The insect toxin

TIC2613 demonstrated activity against CEW, CLW, ECB, FAW, SBL, SAW, SCB, SWCB, TBW, and VBC. Activity was not observed for TIC1425 and TIC2613 when assayed against BCW.

In a separate series of experiments, protein preparations of 5 TIC4472 were assayed using a diet overlay assay against the Lepidopteran insect pests Beet armyworm (BAW, Spodoptera exigua), Pink bollworm (PBW, Pectinophora gossypiella), Cry1Ac resistant Pink bollworm (PBW_Cry1Ac^r, Pectinophora gossypiella), Old world bollworm (OWB, Helicoverpa armigera), Oriental leaf worm (OLW, Spodoptera litura), and Spotted bollworm (SBW, Earias vittella). Table 5 shows the activity observed against each of these Lepidopteran insect pests assayed in the diet overlay bioassay, wherein "+" indicates activity.

TABLE 5

Bioassa	y activity	of TIC4472 against I	_epidoptera	an insect p	ests.
BAW	PBW	PBW_Cry1Ac ^r	OWB	OLW	SBW
+	+	+	+	+	+

Lepidopteran Insect Pest

As can be seen in Table 5 above, TIC4472 demonstrated activity against all of the Lepidopteran insect pests assayed ²⁵ in the diet overlay bioassay, including the Cry1Ac resistant colony of Pink bollworm.

As demonstrated in Tables 3-5, TIC4472, TIC1425, and TIC2613 demonstrate activity across a wide range of Lepidopteran insect pest species.

Example 3

Design of Synthetic Coding Sequences Encoding TIC4472PL and TIC2613PL for Expression in Plant Cells

Synthetic coding sequences were constructed for use in expression of the encoded protein in plants, cloned into a binary plant transformation vectors, and used to transform 40 plant cells. The synthetic sequences were synthesized according to methods generally described in U.S. Pat. No. 5,500,365, avoiding certain inimical problem sequences such as ATTTA and A/T rich plant polyadenylation sequences. The synthetic coding sequences encode 45 TIC4472PL and TIC2613PL proteins which comprise an additional alanine residue immediately following the initiating methionine relative to the TIC4472 and TIC2613 proteins, respectively. The additional alanine residue is incorporated into the synthetic coding sequence to improve 50 expression of the insect toxin protein.

The synthetic coding sequence encoding TIC4472PL (SEQ ID NO:4) is presented herein as SEQ ID NO:3. The synthetic coding sequence encoding TIC2613PL (SEQ ID NO:10) is presented as SEQ ID NO:9. The synthetic coding sequences were used in binary plant transformation vectors to generate transgenic plants expressing the TIC4472PL and TIC2613PL proteins and assayed for activity against Lepidopteran insect pests.

Example 4

Assay of TIC4472PL Activity Against Lepidopteran Pests in Stably Transformed Cotton Plants

Binary plant transformation vectors comprising transgene cassettes designed to express untargeted TIC4472PL pesti-

28

cidal protein were cloned using methods known in the art. The resulting vectors were used to stably transform cotton plants. Tissues were harvested from the transformants and used in insect bioassay against various Lepidopteran insect pests.

Synthetic coding sequences were constructed for use in expression of the encoded protein in plants, cloned into a binary plant transformation vector, and used to transform cotton plant cells. The resulting plant transformation vectors comprised a first transgene cassette for expression of the TIC4472PL pesticidal protein which comprised a constitutive promoter, operably linked 5' to a leader, operably linked 5' to a synthetic coding sequence encoding an untargeted TIC4472PL protein, which was in turn operably linked 5' to a 3' UTR; and a second transgene cassette for the selection of transformed plant cells using spectinomycin selection.

Three binary plant transformation vector constructs were constructed as described above and comprised different 20 constitutive promoters. Two constructs, Construct 1 and Construct 3 also comprised an intron sequence that was operably linked 3' to the leader and 5' to the synthetic coding sequence encoding an untargeted TIC4472PL protein. Construct 2 was intronless. Cotton plants were transformed with the three different binary transformation vectors using an Agrobacterium-mediated transformation method. The transformed cells were induced to form plants by methods known in the art. Bioassays using plant leaf disks were performed analogous to those described in U.S. Pat. No. 8,344,207. A single freshly hatched neonate larvae less than one day old was placed on each leaf disc sample and allowed to feed for approximately four days. A non-transformed cotton plant was used to obtain tissue to be used as a negative control. Multiple transformation R₀ single-copy insertion events from each binary vector were assessed against Cotton bollworm (CBW, Helicoverpa zea), Fall armyworm (FAW, Spodoptera frugiperda), Soybean looper (SBL, Chrysodeixis includens), and Tobacco budworm (TBW, Heliothis virescens). Leaf Damage Rating (LDR) scores ranging from one (1) to four (4) were applied to event. The LDR scores are based upon the percentage of damage observed for each leaf disc. Table 6 below shows the LDR scores and corresponding range of percent damage associated with each score. For LDR scores of one (1), the penetrance of the trait is also assessed. High penetrance (as indicated by "(H)") is defined as greater than fifty (50) percent of the assayed events for each construct having less than or equal to ten (10) percent leaf damage. Low penetrance (as indicated by "(L)") is defined as less than or equal to fifty (50) percent of the assayed events for each construct having less than or equal to ten (10) percent leaf damage. Penetrance is not applied to LDR scores greater than one (1).

TABLE 6

33	Cotton Lea	af Damage Rating (LDR) so percentage damage and l	
	LDR Score	Leaf Damage Rating (LDR) range	Penetrance
60	1	LDR ≤ 10%	High = >50%
	1		of events Low = <50% of events
	2	$10\% < LDR \le 17.5\%$	n/a
	3	$17.5\% \le LDR \le 35\%$	n/a
65	4	LDR > 35%	n/a

The leaf damage rating scores for transformed $R_{\rm o}$ cotton plants expressing TIC4472PL are presented in Table 7 below. The number of events demonstrating the LDR score out of the total number events assayed is shown in parenthesis, followed by the penetrance for those LDR scores of 5 one (1).

TABLE 7

Leaf Damage Rating (LDR) scores, number of events
demonstrating the LDR, and penetrance for transformed
R ₀ cotton plants expressing TIC4472PL.

Construct	CBW	FAW	SBL	TBW
Construct 1	3 (1/25)	2 (2/25)	1 (25/25) H	1 (24/25) H
Construct 2	3 (3/25)	2 (1/25)	1 (24/25) H	1 (16/25) H
Construct 3	4 (23/23)	1 (1/23) L	1 (21/23) H	1 (6/23) L

Transformed R_0 cotton plants expressing TIC4472PL were highly efficacious (defined as having less than or equal to ten percent leaf damage) against SBL and TBW as shown in Table 7. Activity against CBW and FAW was also observed in several events.

 R_1 cotton events were selected from the transformed R_0 cotton plants assayed above and were used in a leaf disc assay against FAW, SBL, and TBW. Table 8 shows the leaf 25 damage rating scores for transformed R_1 cotton plants expressing TIC4472PL.

TABLE 8

Leaf Damage Rating (LDR) scores, number of events demonstrating the LDR, and penetrance for transformed R_1 cotton plants expressing TIC4472PL.

Construct	FAW	SBL	TBW
Construct 1	1 (3/4) H	1 (4/4) H	1 (4/4) H
Construct 2	1 (1/4) L	1 (4/4) H	1 (1/4) L

As can be seen in Table 8 above, the selected events showed high efficacy against FAW, SBL, and TBW. Penetrance was high for Construct 1 transformed events for all three insect pest species. With respect to Construct 2 transformed events, penetrance was high for SBL.

The forgoing demonstrates that transformed cotton plants expressing TIC4472PL protein provide resistance to Fall armyworm (FAW, *Spodoptera frugiperda*), Soybean looper ⁴⁵ (SBL, *Chrysodeixis includens*) and Tobacco budworm (TBW, *Heliothis virescens*).

Example 5

Assay of TIC4472PL and TIC2613PL Activity Against Lepidopteran Pests in Stably Transformed Soybean Plants

A binary plant transformation vector comprising a transgene cassette designed to express a plastid targeted TIC4472PL or TIC2613PL untargeted pesticidal protein was cloned using methods known in the art. The resulting vector was used to stably transform soybean plants. Tissues were harvested from the transformants and used in insect bioassay against various Lepidopteran insect pests.

The synthetic coding sequences designed for plant expression described in Example 3 were cloned into binary plant transformation vector constructs, and used to transform soybean plant cells. Three binary vector constructs were constructed using methods known in the art to express 65 plastid targeted and untargeted TIC4472PL. Construct 4 comprised a first transgene cassette for expression of the

untargeted TIC4472PL which comprised a constitutive promoter, operably linked 5' to a leader, operably linked 5' to a synthetic coding sequence encoding an untargeted TIC4472PL protein, which was in turn operably linked 5' to a 3' UTR and; a second transgene cassette for the selection of transformed plant cells using spectinomycin selection. Construct 5 comprised a first transgene cassette for expression of the targeted TIC4472PL which comprised a constitutive promoter, operably linked 5' to a leader, operably linked 5' to an intron, operably linked 5' to a synthetic coding sequence encoding a plastid targeted TIC4472PL protein, which was in turn operably linked 5' to a 3' UTR and; a second transgene cassette for the selection of transformed plant cells using spectinomycin selection. Construct 6 comprised a first transgene cassette for expression of the untargeted TIC4472PL which comprised a constitutive promoter, operably linked 5' to a leader, operably linked 5' to an intron, operably linked 5' to a synthetic coding sequence encoding an untargeted TIC4472PL protein, which was in turn operably linked 5' to a 3' UTR and; a second transgene cassette for the selection of transformed plant cells using spectinomycin selection.

Two binary vector constructs were constructed using methods known in the art to express untargeted TIC2613PL. Construct 7 comprised a first transgene cassette for expression of the untargeted TIC2613PL which comprised a constitutive promoter, operably linked 5' to a leader, operably linked 5' to a synthetic coding sequence encoding an untargeted TIC2613PL protein, which was in turn operably linked 5' to a 3' UTR and; a second transgene cassette for the selection of transformed plant cells using spectinomycin selection. Construct 8 comprised a first transgene cassette for expression of the untargeted TIC2613PL which comprised a constitutive promoter, operably linked 5' to a leader, operably linked 5' to an intron, operably linked 5' to a synthetic coding sequence encoding an untargeted TIC2613PL protein, which was in turn operably linked 5' to a 3' UTR and; a second transgene cassette for the selection of transformed plant cells using spectinomycin selection.

The resulting binary transformation vector constructs were used to transform soybean cells using an Agrobacterium-mediated transformation method. The transformed soybean cells were induced to form plants by methods known in the art. Bioassays using plant leaf disks were performed analogous to those described in U.S. Pat. No. 8,344,207. A non-transformed soybean plant was used to obtain tissue to be used as a negative control. Multiple R_o soybean transformation events derived from the binary vectors were assessed against Southern armyworm (SAW, Spodoptera eridania), Soybean looper (SBL, Chrysodeixis includens), and Soybean podworm (SPW, Helicoverpa zea). The Leaf Damage Rating (LDR) scores were similar to those as for cotton but differed in the range of percentage damage used to determine the score. The Leaf Damage Rating scores and their corresponding percent leaf damage rating range are presented in Table 9 below.

TABLE 9

LDR Score	Leaf Damage Rating (LDR) range	
1 2	LDR < 20% 20% < LDR < 35%	
3 4	35% < LDR ≤ 70% LDR > 70%	
	LDR Score	Score (LDR) range 1 LDR \(\leq 20\)% 2 20\% < LDR \(\leq 35\)% 3 35\% < LDR \(\leq 70\)%

The leaf damage rating scores for transformed $R_{\rm o}$ cotton plants expressing TIC4472PL and TIC2613PL are presented in Table 10 below. The number of events demonstrating the LDR score out of the total number events assayed is shown in parenthesis.

TABLE 10

Leaf Damage Rating (LDR) scores and number of events demonstrating the LDR for transformed R₀ soybean plants expressing TIC4472PL and TIC2613PL.

Toxin	Construct	SAW	SBL	SPW
TIC4472PL	Construct 4 Construct 5 Construct 6 Construct 7 Construct 8	1 (23/29)	1 (29/29)	3 (8/29)
TIC4472PL		1 (20/29)	1 (28/29)	1 (1/29)
TIC4472PL		1 (10/29)	1 (15/15)	3 (3/15)
TIC2613PL		1 (22/30)	1 (24/30)	3 (1/30)
TIC2613PL		1 (14/25)	1 (22/25)	2 (2/25)

As can be seen in Table 10 above, both expression of both TIC4472PL and TIC2613PL demonstrated high efficacy against SAW and SBL. Activity was lower for both proteins against SPW.

 R_1 soybean events were selected from the transformed R_0 soybean plants expressing TIC4472PL and TIC2613PL assayed above and were used in a leaf disc assay against SAW, SBL, and SPW. A R_1 soybean event expressing TIC4472PL was also assessed against Velvet bean caterpillar (VBW, *Anticarsia gemmatalis*). Table 11 shows the leaf damage rating scores for transformed R_1 soybean plants expressing TIC4472PL and TIC2613PL.

TABLE 11

Leaf Damage Rating (LDR) scores and number of events demonstrating the LDR for transformed R₁ soybean plants expressing TIC4472PL and TIC2613PL.

Toxin	Construct	SAW	SBL	SPW	VBC
TIC4472PL	Construct 5	1 (3/8)	1 (5/8)	3 (2/8)	1 (3/9)
TIC4472PL	Construct 6	1 (1/9)	1 (6/9)	3 (1/9)	
TIC2613PL	Construct 8	2 (2/10)	1 (9/10)	4 (4/10)	

As can be seen in Table 11, a number of R_1 soybean events transformed with TIC4472PL demonstrated high efficacy against SAW and SBL. In addition, several R_1 soybean events transformed using Construct 6 demonstrated high ⁴⁵ efficacy against VBC. R_1 soybean events transformed with TIC2613PL demonstrated high efficacy against SBL.

The forgoing demonstrates that transformed soybean plants expressing TIC4472PL or TIC2613PL provide resistance to Lepidopteran insects, in particular Southern armyworm (*Spodoptera eridania*) and Soybean looper (*Chrysodeixis includens*) and Velvet bean caterpillar (*Anticarsia gemmatalis*).

Example 6

Assay of TIC4472PL and TIC2613PL Activity Against Lepidopteran Pests in Stably Transformed Corn Plants

Binary plant transformation vectors comprising transgene cassettes designed to express both plastid targeted and untargeted TIC4472PL or TIC2613PL pesticidal protein are cloned using methods known in the art. The resulting vectors are used to stably transform corn plants. Tissues are harvested from the transformants and used in insect bioassay against various Lepidopteran insect pests.

The synthetic coding sequence encoding TIC4472PL (SEQ ID NO:3) or TIC2613 (SEQ ID NO:9) protein are cloned into binary transformation vectors. For plastid targeted protein, the synthetic TIC4472PL or TIC2613PL pesticidal protein coding sequence is operably linked in frame with a chloroplast targeting signal peptide encoding sequence. The resulting plant transformation vectors comprise a first transgene cassette for expression of the TIC4472PL or TIC2613PL pesticidal protein which comprise a constitutive promoter, operably linked 5' to a leader, operably linked 5' to an intron, operably linked 5' to a synthetic coding sequence encoding a plastid targeted or untargeted TIC4472PL or TIC2613PL protein, which is in turn operably linked 5' to a 3' UTR; and a second transgene cassette for the selection of transformed plant cells using glyphosate selection.

Corn plants are transformed with the binary transformation vectors described above using an Agrobacterium-mediated transformation method. The transformed cells are induced to form plants by methods known in the art. Bioassays using plant leaf disks are performed analogous to those described in U.S. Pat. No. 8,344,207. A non-transformed corn plant is used to obtain tissue to be used as a negative control. Multiple transformation events from each binary vector are assessed against Beet armyworm (BAW, Spodoptera exigua), Black cutworm (BCW, Agrotis ipsilon), Corn earworm (CEW, Helicoverpa zea), Cotton leaf worm (CLW, Alabama argillacea), European corn borer (ECB, Ostrinia nubilalis), Fall armyworm (FAW, Spodoptera frugiperda), Old world bollworm (OWB, Helicoverpa armigera), Oriental leaf worm (OLW, Spodoptera litura), Pink bollworm (PBW, Pectinophora gossypiella), Soybean looper (SBL, Chrysodeixis includens), Spotted bollworm (SBW, Earias vittella), Southwestern corn borer (SWCB, Diatraea grandiosella), Sugarcane borer (SCB, Diatraea saccharalis), Tobacco budworm (TBW, Heliothis virescens), and Velvet bean caterpillar (VBW, Anticarsia gemmatalis), as well as other Lepidopteran insect pests.

The insect pests are observed for mortality and stunting caused by ingestion of the presented leaf discs expressing TIC4472PL or TIC2613PL and compared to leaf discs derived from non-transformed corn plants.

All of the compositions disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions of this invention have been described in terms of the foregoing illustrative embodiments, it will be apparent to those of skill in the art that variations, changes, modifications, and alterations may be applied to the composition described herein, without departing from the true concept, spirit, and scope of the invention. More specifically, it will be apparent that certain agents that are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope, and concept of the invention as defined by the appended claims.

All publications and published patent documents cited in the specification are incorporated herein by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

SEQUENCE LISTING

```
<160> NUMBER OF SEQ ID NOS: 10
<210> SEQ ID NO 1
<211> LENGTH: 3564
<212> TYPE: DNA
<213> ORGANISM: Bacillus thuringienses
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(3564)
<223> OTHER INFORMATION: Nucleic acid sequence encoding a TIC4472
     pesticidal protein obtained from Bacillus thuringiensis species
      EG10742.
<400> SEQUENCE: 1
atgaataata atattgaaaa ccaatgcgta ccttacaatt gtttaagtaa tcctgaagaa
                                                                      60
gtaattttgg atggagaacg gatatcaact ggtaattcat caattgatat ttctctgtca
                                                                     120
cttgttcaac ttctggtatc taactttgta ccaggcggag gatttttagt agggttaata
                                                                     180
gattttgtat ggggaatagt aggcccttct ccatgggatg catttctagt gcaaattgaa
                                                                     240
caattaattc agcaaagaat agaagcatat gctagggctg cagcaatttc taatttagaa
                                                                     300
ggaataggaa acaatttcaa tatatatgtg gaagcatttc aagaatggga agaagatcct
                                                                     360
aataatccag caaccaggaa tagagtagtt gategettte gtataettga tgggetaett
                                                                     420
gaaagggaca ttccttcgtt tcgaatttct ggatttgaag tccccctttt atccgtttat
                                                                     480
gctcaagcgg ccaatctgca tctagctata ttaagagatt ctgtaatttt tggagaaaga
                                                                     540
tggggattga caacaacaaa tgtcaatgaa aactataata gacaaatcag gcatattgat
                                                                     600
gaatatgctg atcactgtgc aaatacgtat aatcggggat taaataattt accgaaatct
                                                                     660
acgtatcaag attggataac atataatcga ttacggagag aattaacatt gactgtatta
                                                                     720
gatatcgctg ctttctttcc aaactatgac aataggcggt atccaattca gccagttggt
                                                                     780
caactaacaa gggaagttta tacggaccca ttaattactt ttaatcccca gttacagtct
                                                                     840
gtageteaat tacetacttt taaegttatg gaaageaaeg caattagaaa teeteatttg
                                                                     900
tttgatatat tgaataatct tacaattttt acggattggt ttagtgttgg acgcaacttt
tattggggag gacatcgagt aacttctaac tatataggag gaggcaacat aacatctcct
                                                                    1020
atatatggaa gagaggcgaa ccaggagcct ccaagatctt ttacttttaa tggacctgtt
                                                                    1080
tttaggactt tatcaaatcc tactttacga ttattacagc aaccttggcc agcaccacca
                                                                    1140
                                                                    1200
tttaatttac gtggtgttga aggagtagaa ttttctacac ctacaaatag ctttacgtat
cgaggaagag gtacagttga ttctttaacc gaattaccgc ctgaggataa tagtgtgcta
                                                                    1260
cctcgcgagg gatatagtca tcgtttatgt catgcaactt ttgttcaaag atctggaaca
                                                                    1320
ccatttttaa caacgggtgt agtattttct tggacgcatc gcagtgcaac tcttacaaat
                                                                    1380
acaattgatc cagacaaaat tactcagata cctttagtga aaggatttag agtttggagt
                                                                    1440
                                                                    1500
ggcgcctctg tcgttacagg accaggtttt acaggagggg atatccttcg aagaactaac
tttggggatt ttgtatctat gcaagttaat attaattcac caataacaca aagataccgt
                                                                    1560
ttaagatttc gttatgcttc cagtagagat gcacgactta cagtagcgac aggagcagca
                                                                    1620
aacacaggag ttggagggca aattagtgtg gacatggctc ttcagaaaac tatggaaatt
                                                                    1680
ggagagaget taacatetag aacatttaga tatacegatt ttagtaatee ttttteattt
                                                                    1740
agagctaatc cagatataat tcgtataaat gaacaaccac tattcggtgc aggctctatt
                                                                    1800
agtagtggtg aactttatat agataaaatt gaaattattc tagcagatgc aacatttgaa
                                                                    1860
gcagaatatg atttggaaag agcacagaag gcggtgaatg cgctgtttac ttctacaaac
                                                                    1920
```

-continued

```
caaagaggat taaaaacaga tgtgacggat tatcatattg atcaagtatc caatttagtt
                                                                    1980
gagtgtttat cggatgaatt ttgtctggat gaaaaacgag aattgtccga gaaagtcaaa
                                                                    2040
catgcgaagc gactcagtga tgagcgaaat ttactccagg atcgaaattt cacatccatt
                                                                    2100
aatgggcaac tagaccgtgg ctggagagga agtacggata ttaccatcca aggaggagat
                                                                    2160
gacgtattca aagagaatta cgtcacacta ctgggtacct ttgatgagtg ctatccaacg
                                                                    2220
                                                                    2280
tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct atacccgcta tgaattaaga
gggtatatcg aggatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacac
                                                                    2340
gaaacagtaa atgtgccagg tacaggttcc ttatggccgc tttcagtcga aagtccaatc
                                                                    2400
ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat
                                                                    2460
tgttcctgca gagacgggaa aacatgtgca catcattctc atcatttctc cttggacatt
                                                                    2520
gatgttggat gtacagactt aaatgaagat ctaggtgtat gggtgatatt caagattaag
                                                                    2580
acgcaagatg gtcatgcaag actaggaaat ctagagtttc tcgaagagaa accattagta
                                                                    2640
                                                                    2700
qqaqaaqcqt taqctcqtqt qaaaaqaqcq qaqaaaaaat qqaqaqacaa acqcqaaaaa
                                                                    2760
ttqcaattaq aaacaaatat cqtttacaaa qaqqcaaaaq aaqctqtqqa tqctttattt
qtaaactctc aatatqataq attacaaqtq qatacqaaca ttqccatqat tcatqcqqca
                                                                    2820
gataaacgeg tteatagaat eegagaageg tatetteeag agttatetgt gatteegggt
                                                                    2880
gtcaatgcgg ctattttcga agaattagaa gggtgtgttt tcactgcatt ctccctatat
                                                                    2940
gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg
                                                                    3000
aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa
                                                                    3060
gcagaagtgt cacaagaggt tegtgtetgt ceaggtegtg getatateet tegtgttaca
                                                                    3120
gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca
                                                                    3180
gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg
                                                                    3240
tgtaatgatt atcctgcaaa tcaagaagaa tacgggggtg cgtacacttc tcgtaatcgt
                                                                    3300
ggatatgacg aaacttatgg aagcaattct tccgtatcag ctgattatgc gtcagtttat
                                                                    3360
gaagaaaaag cgtatacaga tggacgaaga gacaatccat gtgaatttaa cagagggtat
                                                                    3420
ggggattata cgccactacc agctggctat gtaacaaaag aattagaata cttcccagaa
                                                                    3480
accgataagg tatggattga gattggagaa acggaaggaa cattcatcgt ggacagtgtg
                                                                    3540
gaattactcc ttatggagga atag
                                                                    3564
```

```
<210> SEQ ID NO 2
```

<400> SEQUENCE: 2

Met Asn Asn Asn Ile Glu Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser 1 $$ 10 $$ 15

Asn Pro Glu Glu Val Ile Leu Asp Gly Glu Arg Ile Ser Thr Gly Asn 20 25 30

Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Leu Leu Val Ser Asn 35 40 45

<211> LENGTH: 1187

<212> TYPE: PRT

<213 > ORGANISM: Bacillus thuringienses

<220> FEATURE:

<221> NAME/KEY: MISC_FEATURE

<222> LOCATION: (1)..(1187)

<223> OTHER INFORMATION: The amino acid sequence of the TIC4472 pesticidal protein.

-continued

Phe	Val 50	Pro	Gly	Gly	Gly	Phe 55	Leu	Val	Gly	Leu	Ile 60	Asp	Phe	Val	Trp
Gly 65	Ile	Val	Gly	Pro	Ser 70	Pro	Trp	Asp	Ala	Phe 75	Leu	Val	Gln	Ile	Glu 80
Gln	Leu	Ile	Gln	Gln 85	Arg	Ile	Glu	Ala	Tyr 90	Ala	Arg	Ala	Ala	Ala 95	Ile
Ser	Asn	Leu	Glu 100	Gly	Ile	Gly	Asn	Asn 105	Phe	Asn	Ile	Tyr	Val 110	Glu	Ala
Phe	Gln	Glu 115	Trp	Glu	Glu	Asp	Pro 120	Asn	Asn	Pro	Ala	Thr 125	Arg	Asn	Arg
Val	Val 130	Asp	Arg	Phe	Arg	Ile 135	Leu	Asp	Gly	Leu	Leu 140	Glu	Arg	Asp	Ile
Pro 145	Ser	Phe	Arg	Ile	Ser 150	Gly	Phe	Glu	Val	Pro 155	Leu	Leu	Ser	Val	Tyr 160
Ala	Gln	Ala	Ala	Asn 165	Leu	His	Leu	Ala	Ile 170	Leu	Arg	Asp	Ser	Val 175	Ile
Phe	Gly	Glu	Arg 180	Trp	Gly	Leu	Thr	Thr 185	Thr	Asn	Val	Asn	Glu 190	Asn	Tyr
Asn	Arg	Gln 195	Ile	Arg	His	Ile	Asp 200	Glu	Tyr	Ala	Asp	His 205	Cya	Ala	Asn
Thr	Tyr 210	Asn	Arg	Gly	Leu	Asn 215	Asn	Leu	Pro	Lys	Ser 220	Thr	Tyr	Gln	Asp
Trp 225	Ile	Thr	Tyr	Asn	Arg 230	Leu	Arg	Arg	Glu	Leu 235	Thr	Leu	Thr	Val	Leu 240
Asp	Ile	Ala	Ala	Phe 245	Phe	Pro	Asn	Tyr	Asp 250	Asn	Arg	Arg	Tyr	Pro 255	Ile
Gln	Pro	Val	Gly 260	Gln	Leu	Thr	Arg	Glu 265	Val	Tyr	Thr	Asp	Pro 270	Leu	Ile
Thr	Phe	Asn 275	Pro	Gln	Leu	Gln	Ser 280	Val	Ala	Gln	Leu	Pro 285	Thr	Phe	Asn
Val	Met 290	Glu	Ser	Asn	Ala	Ile 295	Arg	Asn	Pro	His	Leu 300	Phe	Asp	Ile	Leu
Asn 305	Asn	Leu	Thr	Ile	Phe 310	Thr	Asp	Trp	Phe	Ser 315	Val	Gly	Arg	Asn	Phe 320
Tyr	Trp	Gly	Gly	His 325	Arg	Val	Thr	Ser	Asn 330	Tyr	Ile	Gly	Gly	Gly 335	Asn
Ile	Thr	Ser	Pro 340		Tyr	Gly			Ala		Gln	Glu	Pro 350	Pro	Arg
Ser	Phe	Thr 355	Phe	Asn	Gly	Pro	Val 360	Phe	Arg	Thr	Leu	Ser 365	Asn	Pro	Thr
Leu	Arg 370	Leu	Leu	Gln	Gln	Pro 375	Trp	Pro	Ala	Pro	Pro 380	Phe	Asn	Leu	Arg
Gly 385	Val	Glu	Gly	Val	Glu 390	Phe	Ser	Thr	Pro	Thr 395	Asn	Ser	Phe	Thr	Tyr 400
Arg	Gly	Arg	Gly	Thr 405	Val	Asp	Ser	Leu	Thr 410	Glu	Leu	Pro	Pro	Glu 415	Asp
Asn	Ser	Val	Leu 420	Pro	Arg	Glu	Gly	Tyr 425	Ser	His	Arg	Leu	Сув 430	His	Ala
Thr	Phe	Val 435	Gln	Arg	Ser	Gly	Thr 440	Pro	Phe	Leu	Thr	Thr 445	Gly	Val	Val
Phe	Ser 450	Trp	Thr	His	Arg	Ser 455	Ala	Thr	Leu	Thr	Asn 460	Thr	Ile	Asp	Pro
Asp	ГХа	Ile	Thr	Gln	Ile	Pro	Leu	Val	Lys	Gly	Phe	Arg	Val	Trp	Ser

-continued

Gly Ala Ser Val Val 485 Thr Gly Pro Gly Phe 490 Thr Gly Asp Lev 490 Thr Gly Asp Met Gly Met Gly Met Gly Met Met Asp Phe Asp Asp Phe Asp Asp	n r l e o n
Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser Ser Arg Asp Ala Arg Leu Thr Val Ala Ala Ala Asp Ile Ala Asp Ile I	r l e o n
Arg Asp Ala Arg Leu Thr S35 520 525 Arg Asp Ala Arg Leu Thr S35 Ala Thr Gly Ala Ala Ala Asn Thr Gly Valor S40 Gly Gly Gln Ile Ser Val Asp Met S55 Ala Leu Gln Lys Thr Met Glu Ile S66 Gly Glu Ser Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser Asp S75 Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Arg Ile Asn Glu Glu S80 Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile Asp	l e o n
Gly Glu Glu Ile Ser Val Asp Met Ala Leu Gln Lys Thr Met Glu Ile 566 Gly Glu Ser Leu Thr 565	e o n n
555 566 567 568 569 569 569 569 569 569 569 569 569 569	n n
Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Arg Ile Asn Glu Glu Sso	n p
580 585 590 Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile Asp	p p
	p
Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu Ala Glu Tyr Asp 610 620	า
Leu Glu Arg Ala Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Thr Ass 625 630 630 635	
Gln Arg Gly Leu Lys Thr Asp Val Thr Asp Tyr His Ile Asp Gln Va 645 650 655	L
Ser Asn Leu Val Glu Cys Leu Ser Asp Glu Phe Cys Leu Asp Glu Ly: 660 665 670	3
Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu 675 680 685	1
Arg Asn Leu Leu Gln Asp Arg Asn Phe Thr Ser Ile Asn Gly Gln Let 690 695 700	ı
Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln Gly Gly Asp 705 710 715 726	
Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Leu Gly Thr Phe Asp Glu 725 730 735	1
Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys 740 745 750	3
Ala Tyr Thr Arg Tyr Glu Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp 755 760 765	
Leu Glu Ile Tyr Leu Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asi 770 775 780	ı
Val Pro Gly Thr Gly Ser Leu Trp Pro Leu Ser Val Glu Ser Pro Ile 785 790 795 800	
Gly Lys Cys Gly Glu Pro Asn Arg Cys Thr Pro His Leu Glu Trp Asn 805 810 815	n
Pro Asn Leu Asp Cys Ser Cys Arg Asp Gly Lys Thr Cys Ala His His 820 825 830	3
Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Ass 835 840 845	n
Glu Asp Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly 850 860	Y
His Ala Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Va. 865 870 875 886	
Gly Glu Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg As 885 890 895	Э.

-continued

Lys Arg Glu Lys Leu Gln Leu Glu Thr Asn Ile Val Tyr Lys Glu Ala 905 Lys Glu Ala Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu 920 Gln Val Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Cys Val Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Glu Glu Gln Asn Asp His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val 1015 Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg 1030 1035 Val Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile 1040 1045 His Gly Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys 1055 1060 1065 Val Glu Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp 1075 Tyr Pro Ala As
n Glu Glu Tyr Gly Gly Ala Tyr Thr Ser Arg 1090 1095 Asn Arg Gly Tyr Asp Glu Thr Tyr Gly Ser Asn Ser Ser Val Ser 1105 Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys Ala Tyr Thr Asp Gly 1115 1120 Arg Arg Asp Asn Pro Cys Glu Phe Asn Arg Gly Tyr Gly Asp Tyr 1135 Thr Pro Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe 1150 Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu 1180 <210> SEQ ID NO 3 <211> LENGTH: 3567 <212> TYPE: DNA <213 > ORGANISM: Artificial <220> FEATURE: <223> OTHER INFORMATION: A synthetic coding sequence encoding a TIC4472PL pesticidal protein designed for expression in a plant cell wherein an additional alanine codon is inserted immediately following the initiating methionine codon. <400> SEQUENCE: 3 atggctaata acaacatcga gaaccagtgc gtgccctaca actgcctttc gaacccggag 60 gaagtgatcc tggacggcga aaggatctcg accgggaata gcagcatcga catctcgctt tegetegtge agettetagt eagtaactte gtteegggeg gagggtttet egtgggeett 180 attgacttcg tttggggcat cgtgggccca tctccttggg acgcattcct cgtgcagatc

-continued

44

gagcaactga	tccagcagcg	tatcgaggcg	tacgctaggg	ctgccgctat	ctccaacctg	300
gagggcatcg	gcaacaactt	caacatctac	gttgaagcct	tccaagaatg	ggaggaagat	360
cctaataacc	cagctacgcg	gaacagagtg	gtggatcgct	ttagaatcct	cgacggcctc	420
ctggaaaggg	acatcccgag	cttccgtatt	teeggetteg	aggtgccgct	gctgagcgtg	480
tacgcgcaag	cggccaatct	gcacctggcg	attctccggg	actctgtgat	cttcggcgag	540
cggtggggcc	tcaccaccac	taacgtgaac	gagaactaca	accgccagat	ccgccacatc	600
gacgagtacg	cggaccactg	cgccaacaca	tacaatcgcg	ggctgaacaa	cctccctaag	660
agcacttacc	aagattggat	cacctacaac	aggctccgcc	gggaactcac	tctcacagtc	720
ctcgacatcg	ctgccttctt	cccgaactac	gacaaccgcc	gctacccgat	tcagccagtc	780
ggccagctca	cccgtgaggt	gtacaccgat	ccactgataa	ctttcaatcc	gcagctccag	840
tctgtcgcac	agttgcccac	cttcaacgtc	atggaaagca	acgccatcag	gaacccacac	900
ttgttcgaca	tccttaacaa	cctgactatc	ttcaccgact	ggttcagcgt	cggacggaac	960
ttctactggg	gcggacaccg	cgtcacctca	aactacatcg	gcggcggcaa	cattacttcg	1020
cccatctacg	gccgggaggc	gaatcaggag	ccgccacgca	gctttacatt	caacggtcct	1080
gtgtttcgca	cgttatcgaa	cccgacactc	aggctgctcc	agcagccctg	gcctgcgccg	1140
ccgtttaatt	tgcgcggcgt	cgaaggcgtc	gagttcagta	cgccgaccaa	cagcttcacc	1200
tatcgcggac	gcgggactgt	tgactccctg	acagagctgc	cgccggagga	caactcggtt	1260
ctgccgcgtg	agggctacag	ccataggctt	tgtcacgcga	cctttgtgca	gcgatccggg	1320
acaccgttcc	ttacaaccgg	cgtggtgttc	tcctggacac	accgcagtgc	aactctgacg	1380
aacacgattg	acccagacaa	gatcacgcag	atcccgttag	tgaagggctt	ccgggtttgg	1440
tctggtgcct	ctgtagtcac	tgggcctggc	tttacgggtg	gcgacatcct	ccgtcgcacg	1500
aactttggcg	acttcgtgtc	catgcaagtg	aacattaaca	gccctattac	gcaacgctac	1560
cggctgaggt	tcagatacgc	ttcctcgcgg	gacgcccgtc	ttacggtggc	gacgggcgca	1620
gcgaacactg	gagttggcgg	ccaaatctcc	gtggacatgg	ctttgcagaa	gactatggag	1680
atcggtgagt	ctctcacatc	tegeaegtte	cgctacacgg	atttctccaa	ccctttctcc	1740
ttccgcgcca	atccggacat	cattcggatc	aacgaacagc	cgctcttcgg	cgcgggctcc	1800
atctcatccg	gtgagcttta	cattgataag	attgagataa	ttctggccga	cgcgaccttt	1860
gaggcagagt	atgatctgga	gcgcgcacag	aaggccgtga	acgcgctgtt	tacgtccacg	1920
aaccagcgcg	ggctcaagac	agacgtcaca	gactaccaca	tcgaccaagt	ctccaacctt	1980
gtcgagtgtc	tctccgacga	gttctgcctg	gacgagaagc	gggagcttag	tgagaaggtg	2040
aagcacgcaa	agcgcctgtc	tgacgagcgg	aaccttctac	aagaccgtaa	cttcacctcc	2100
attaacgggc	agctagaccg	tggctggcgc	gggtccaccg	acatcactat	ccaaggtggc	2160
gacgacgtct	tcaaggagaa	ctacgtgacg	ctgctcggca	cctttgacga	gtgctacccg	2220
acatacctct	atcagaagat	tgacgagtct	aagctcaagg	cttacacccg	ttacgagctg	2280
cgtggctaca	tcgaggactc	ccaggatctg	gaaatctatc	tcatcagata	caacgcgaag	2340
cacgagacag	tcaacgtacc	tgggacaggc	tetetetgge	ctctgtctgt	ggagagtccc	2400
atcggcaaat	gtggcgagcc	gaacagatgt	actccgcacc	tggagtggaa	tcccaacttg	2460
	gccgcgacgg					2520
	gctgcacgga					2580
	acggacacgc					2640
		3 5 - 3 - 3 - 3 - 3	Jungage		2333	_010

				-contir	nued	
gtcggcgaag	cactggcgcg	ggtcaagagg	gccgagaaga	agtggcggga	caaacgggag	2700
aaacttcaac	tcgaaacgaa	catcgtttac	aaggaggcaa	aggaggccgt	ggacgcactg	2760
ttcgtgaact	cgcagtacga	ccgcctccaa	gtggacacca	acatcgccat	gatccatgca	2820
geggaeaage	gcgtgcaccg	aatcagggaa	gcgtacttgc	ccgagttgtc	cgttatccct	2880
ggcgtgaacg	ctgccatctt	cgaggaactg	gagggctgcg	tgttcaccgc	attctccctg	2940
tacgacgcac	gaaacgtcat	caagaatggc	gacttcaaca	acggcctgag	ctgctggaac	3000
gtgaagggcc	acgtggagga	gcagaacgac	caccgctccg	tgttagtggt	cccggaatgg	3060
gaagccgaag	tgagccagga	ggtcagggtg	tgtcccggtc	gcggttacat	cctccgcgtg	3120
accgcctaca	aggagggcta	tggcgagggc	tgcgtgacga	tacacggtat	cgagaacaac	3180
accgatgagc	ttaagttctc	gaactgcgtg	gaggaggagg	tgtacccgaa	taacacagtg	3240
acgtgcaatg	actacccggc	caaccaggag	gagtacggcg	gtgcctacac	gagccgaaac	3300
cgtggctatg	acgaaactta	cggctcgaac	agcagcgtgt	ctgcggatta	tgccagtgtg	3360
tacgaggaga	aggcgtacac	ggacgggcgg	cgcgacaacc	cttgcgagtt	caatagaggc	3420
tatggcgact	acacgccgct	gcccgccggt	tatgtgacga	aggagttgga	atacttccca	3480
gagacggaca	aggtgtggat	cgagattggc	gagaccgagg	gcacgttcat	tgtggacagc	3540
gttgagctgc	tactgatgga	ggagtga				3567
<220> FEATU <223> OTHER a syn cell	TH: 1188 : PRT NISM: Artif: URE: R INFORMATIOnthetic cod: (SEQ ID NO	ON: The aming sequence	e designed i erein an ado	for express: ditional ala	C4472PL encodion in a plan anine amino a methionine.	nt
<400> SEQUI	ENCE: 4					
Mat Ala Acr	n Aen Aen T	le Clu Aen (Tin Cvc Val	Dro Tur Aci	o Cva Leu	

Met Ala Asn Asn Asn Ile Glu Asn Gln Cys Val Pro Tyr Asn Cys Leu

Ser Asn Pro Glu Glu Val Ile Leu Asp Gly Glu Arg Ile Ser Thr Gly $20 \\ 25 \\ 30$

Asn Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Leu Leu Val Ser

Asn Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val

Trp Gly Ile Val Gly Pro Ser Pro Trp Asp Ala Phe Leu Val Gln Ile 65 70 70 75 75 80

Glu Gl
n Leu Ile Gl
n Gln Arg Ile Glu Ala Tyr Ala Arg Ala Ala Ala 85 90 95

Ile Ser Asn Leu Glu Gly Ile Gly Asn Asn Phe Asn Ile Tyr Val Glu 105

Ala Phe Gln Glu Trp Glu Glu Asp Pro Asn Asn Pro Ala Thr Arg Asn

Arg Val Val Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp 135

Ile Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val 150 155

Tyr Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val 170

-continued
-concinued

_															
Ile	Phe	Gly	Glu 180	Arg	Trp	Gly	Leu	Thr 185	Thr	Thr	Asn	Val	Asn 190	Glu	Asn
Tyr	Asn	Arg 195	Gln	Ile	Arg	His	Ile 200	Asp	Glu	Tyr	Ala	Asp 205	His	Cys	Ala
Asn	Thr 210	Tyr	Asn	Arg	Gly	Leu 215	Asn	Asn	Leu	Pro	Lys 220	Ser	Thr	Tyr	Gln
Asp 225	Trp	Ile	Thr	Tyr	Asn 230	Arg	Leu	Arg	Arg	Glu 235	Leu	Thr	Leu	Thr	Val 240
Leu	Asp	Ile	Ala	Ala 245	Phe	Phe	Pro	Asn	Tyr 250	Asp	Asn	Arg	Arg	Tyr 255	Pro
Ile	Gln	Pro	Val 260	Gly	Gln	Leu	Thr	Arg 265	Glu	Val	Tyr	Thr	Asp 270	Pro	Leu
Ile	Thr	Phe 275	Asn	Pro	Gln	Leu	Gln 280	Ser	Val	Ala	Gln	Leu 285	Pro	Thr	Phe
Asn	Val 290	Met	Glu	Ser	Asn	Ala 295	Ile	Arg	Asn	Pro	His 300	Leu	Phe	Asp	Ile
Leu 305	Asn	Asn	Leu	Thr	Ile 310	Phe	Thr	Asp	Trp	Phe 315	Ser	Val	Gly	Arg	Asn 320
Phe	Tyr	Trp	Gly	Gly 325	His	Arg	Val	Thr	Ser 330	Asn	Tyr	Ile	Gly	Gly 335	Gly
Asn	Ile	Thr	Ser 340	Pro	Ile	Tyr	Gly	Arg 345	Glu	Ala	Asn	Gln	Glu 350	Pro	Pro
Arg	Ser	Phe 355	Thr	Phe	Asn	Gly	Pro 360	Val	Phe	Arg	Thr	Leu 365	Ser	Asn	Pro
Thr	Leu 370	Arg	Leu	Leu	Gln	Gln 375	Pro	Trp	Pro	Ala	Pro 380	Pro	Phe	Asn	Leu
Arg 385	Gly	Val	Glu	Gly	Val 390	Glu	Phe	Ser	Thr	Pro 395	Thr	Asn	Ser	Phe	Thr 400
Tyr	Arg	Gly	Arg	Gly 405	Thr	Val	Asp	Ser	Leu 410	Thr	Glu	Leu	Pro	Pro 415	Glu
Asp	Asn	Ser	Val 420	Leu	Pro	Arg	Glu	Gly 425	Tyr	Ser	His	Arg	Leu 430	Cys	His
Ala	Thr	Phe 435	Val	Gln	Arg	Ser	Gly 440	Thr	Pro	Phe	Leu	Thr 445	Thr	Gly	Val
Val	Phe 450	Ser	Trp	Thr	His	Arg 455	Ser	Ala	Thr	Leu	Thr 460	Asn	Thr	Ile	Asp
Pro 465	Asp	Lys	Ile		Gln 470		Pro	Leu		Lys 475	Gly	Phe	Arg		Trp 480
Ser	Gly	Ala	Ser	Val 485	Val	Thr	Gly	Pro	Gly 490	Phe	Thr	Gly	Gly	Asp 495	Ile
Leu	Arg	Arg	Thr 500	Asn	Phe	Gly	Asp	Phe 505	Val	Ser	Met	Gln	Val 510	Asn	Ile
Asn	Ser	Pro 515	Ile	Thr	Gln	Arg	Tyr 520	Arg	Leu	Arg	Phe	Arg 525	Tyr	Ala	Ser
Ser	Arg 530	Asp	Ala	Arg	Leu	Thr 535	Val	Ala	Thr	Gly	Ala 540	Ala	Asn	Thr	Gly
Val 545	Gly	Gly	Gln	Ile	Ser 550	Val	Asp	Met	Ala	Leu 555	Gln	Lys	Thr	Met	Glu 560
Ile	Gly	Glu	Ser	Leu 565	Thr	Ser	Arg	Thr	Phe 570	Arg	Tyr	Thr	Asp	Phe 575	Ser
Asn	Pro	Phe	Ser 580	Phe	Arg	Ala	Asn	Pro 585	Asp	Ile	Ile	Arg	Ile 590	Asn	Glu
Gln	Pro	Leu	Phe	Gly	Ala	Gly	Ser	Ile	Ser	Ser	Gly	Glu	Leu	Tyr	Ile

-continued	
-continued	

		595					600					605			
Asp	Lys 610	Ile	Glu	Ile	Ile	Leu 615	Ala	Asp	Ala	Thr	Phe 620	Glu	Ala	Glu	Tyr
Asp 625	Leu	Glu	Arg	Ala	Gln 630	Lys	Ala	Val	Asn	Ala 635	Leu	Phe	Thr	Ser	Thr 640
Asn	Gln	Arg	Gly	Leu 645	Lys	Thr	Asp	Val	Thr 650	Asp	Tyr	His	Ile	Asp 655	Gln
Val	Ser	Asn	Leu 660	Val	Glu	CAa	Leu	Ser 665	Asp	Glu	Phe	СЛа	Leu 670	Asp	Glu
Lys	Arg	Glu 675	Leu	Ser	Glu	Lys	Val 680	Lys	His	Ala	Lys	Arg 685	Leu	Ser	Asp
Glu	Arg 690	Asn	Leu	Leu	Gln	Asp 695	Arg	Asn	Phe	Thr	Ser 700	Ile	Asn	Gly	Gln
Leu 705	Asp	Arg	Gly	Trp	Arg 710	Gly	Ser	Thr	Asp	Ile 715	Thr	Ile	Gln	Gly	Gly 720
Asp	Asp	Val	Phe	Lys 725	Glu	Asn	Tyr	Val	Thr 730	Leu	Leu	Gly	Thr	Phe 735	Asp
Glu	Cys	Tyr	Pro 740	Thr	Tyr	Leu	Tyr	Gln 745	Lys	Ile	Asp	Glu	Ser 750	ГÀа	Leu
rys	Ala	Tyr 755	Thr	Arg	Tyr	Glu	Leu 760	Arg	Gly	Tyr	Ile	Glu 765	Asp	Ser	Gln
Asp	Leu 770	Glu	Ile	Tyr	Leu	Ile 775	Arg	Tyr	Asn	Ala	Lys 780	His	Glu	Thr	Val
Asn 785	Val	Pro	Gly	Thr	Gly 790	Ser	Leu	Trp	Pro	Leu 795	Ser	Val	Glu	Ser	Pro 800
Ile	Gly	Lys	Cys	Gly 805	Glu	Pro	Asn	Arg	Cys	Thr	Pro	His	Leu	Glu 815	Trp
Asn	Pro	Asn	Leu 820	Asp	Сув	Ser	Сув	Arg 825	Asp	Gly	Lys	Thr	630	Ala	His
His	Ser	His 835	His	Phe	Ser	Leu	Asp 840	Ile	Asp	Val	Gly	Cys 845	Thr	Asp	Leu
Asn	Glu 850	Asp	Leu	Gly	Val	Trp 855	Val	Ile	Phe	Lys	Ile 860	Lys	Thr	Gln	Asp
Gly 865	His	Ala	Arg	Leu	Gly 870	Asn	Leu	Glu	Phe	Leu 875	Glu	Glu	Lys	Pro	Leu 880
Val	Gly	Glu	Ala	Leu 885	Ala	Arg	Val	Lys	Arg 890	Ala	Glu	Lys	Lys	Trp 895	Arg
Asp	Lys	Arg	Glu 900	Lys	Leu	Gln	Leu	Glu 905	Thr	Asn	Ile	Val	Tyr 910	Lys	Glu
Ala	Lys	Glu 915	Ala	Val	Asp	Ala	Leu 920	Phe	Val	Asn	Ser	Gln 925	Tyr	Asp	Arg
Leu	Gln 930	Val	Asp	Thr	Asn	Ile 935	Ala	Met	Ile	His	Ala 940	Ala	Asp	Lys	Arg
Val 945	His	Arg	Ile	Arg	Glu 950	Ala	Tyr	Leu	Pro	Glu 955	Leu	Ser	Val	Ile	Pro 960
Gly	Val	Asn	Ala	Ala 965	Ile	Phe	Glu	Glu	Leu 970	Glu	Gly	CÀa	Val	Phe 975	Thr
Ala	Phe	Ser	Leu 980	Tyr	Asp	Ala	Arg	Asn 985	Val	Ile	Lys	Asn	Gly 990	Asp	Phe
Asn	Asn	Gly 995	Leu	Ser	Сув	Trp	Asn 1000		L Lys	s Gl	y Hi:	s Vai		lu G	lu Gln
Asn	Asp 1010		s Arq	g Se:	r Vai	l Le: 10:		al Va	al Pi	ro Gi		rp (Glu <i>l</i>	Ala (Glu

-continued

Val	Ser 1025	Gln	Glu	Val	Arg	Val 1030		Pro	Gly	Arg	Gly 1035	Tyr	Ile	Leu	
Arg	Val 1040	Thr	Ala	Tyr	Lys	Glu 1045		Tyr	Gly	Glu	Gly 1050	Сув	Val	Thr	
Ile	His 1055	Gly	Ile	Glu	Asn	Asn 1060		Asp	Glu	Leu	Lys 1065	Phe	Ser	Asn	
Сув	Val 1070	Glu	Glu	Glu	Val	Tyr 1075	Pro	Asn	Asn	Thr	Val 1080	Thr	Сув	Asn	
Asp	Tyr 1085	Pro	Ala	Asn	Gln	Glu 1090	Glu	Tyr	Gly	Gly	Ala 1095	Tyr	Thr	Ser	
Arg	Asn 1100	Arg	Gly	Tyr	Asp	Glu 1105	Thr	Tyr	Gly	Ser	Asn 1110	Ser	Ser	Val	
Ser	Ala 1115	Asp	Tyr	Ala	Ser	Val 1120		Glu	Glu	Lys	Ala 1125	Tyr	Thr	Asp	
Gly	Arg 1130		Asp	Asn	Pro	Cys 1135	Glu	Phe	Asn	Arg	Gly 1140	Tyr	Gly	Asp	
Tyr	Thr 1145	Pro	Leu	Pro	Ala	Gly 1150		Val	Thr	Lys	Glu 1155	Leu	Glu	Tyr	
Phe	Pro 1160	Glu	Thr	Asp	Lys	Val 1165		Ile	Glu	Ile	Gly 1170	Glu	Thr	Glu	
Gly	Thr 1175	Phe	Ile	Val	Asp	Ser 1180	Val	Glu	Leu	Leu	Leu 1185	Met	Glu	Glu	
<212 <213 <220 <221 <222 <223	0> FEA L> NAI 2> LOO 3> OTI pes	PE: 1 GANI: ATURI ME/KI CATIC HER : stic:	DNA SM: F E: EY: T ON: INFOR idal	Bacil misc_ (1) RMATI prot	feat (356	Mucle	∍ic a	acid	seqı					TIC1425 sis spec	cies
atqa	aataai	ca at	tatto	qaaaa	ı cca	aatqc	qta (cctta	acaat	t qt	tttaad	qtaa	tcci	tqaaqaa	60
gtaa	atttt	gg at	tggag	gaaco	g gat	atcaa	act o	ggtaa	attca	at ca	aattga	atat	ttci	tctgtca	120
ctto	gttcaa	ac ti	tctg	gtato	taa	acttt	gta d	ccago	gegga	ag ga	atttti	agt	agg	gttaata	180
gatt	ttgt	at g	gggaa	atagt	agg	gecett	cct (ccato	gggat	g ca	atttci	agt	gcaa	aattgaa	240
caat	taati	cc a	gcaaa	agaat	aga	aagcat	at q	gctag	gggct	g ca	agcaat	ttc	taat	ttagaa	300
ggaa	atagga	aa a	caatt	tcaa	a tat	atat	gtg g	gaago	catt	cc aa	agaat	ggga	agaa	agatcct	360
aata	aatcca	ag c	aacca	aggaa	a tag	gagtaç	gtt (gatco	gcttt	c gt	tatact	tga	tgg	gctactt	420
gaaa	aggga	ca t	tcctt	cgtt	teg	gaatti	ct q	ggatt	tgaa	ag to	cccct	ttt	atc	egtttat	480
gcto	caagc	gg c	caato	ctgca	a tct	agcta	ata t	taaç	gagat	t ct	tgtaat	ttt	tgg	agaaaga	540
tggg	ggatt	ga c	aacaa	acaaa	ı tgt	caat	gaa a	aacta	ataat	a ga	acaaat	cag	gcat	tattgat	600
gaat	atgci	g a	tcact	gtgo	c aaa	atacgt	at a	aatco	9999	at ta	aaataa	attt	acc	gaaatct	660
acgt	atca	ag at	ttgga	ataac	ata	ataato	ega t	tace	ggaga	ag aa	attaad	catt	gact	tgtatta	720
gata	ategei	eg et	tttct	ttcc	c aaa	actato	gac a	aatag	ggcgg	gt at	tccaat	tca	gcca	agttggt	780
caad	ctaaca	aa g	ggaaq	gttta	a tao	eggaco	cca t	taat	tact	t ti	taatc	cca	gtta	acagtct	840
gtag	gctcaa	at ta	accta	acttt	t taa	acgtta	atg 🤉	gaaag	gcaad	eg ca	aatta	gaaa	tcci	tcatttg	900
														caacttt	960
								55-	- 3:		5-5	رر	- 0		

-continued

tetatggaa gagagggaa cagaggget caagatett thatttaa tggacctgtt 1140 tttaatttac gtggtgttga aggatagaa tttetacage aacettggee agcaccacca 1140 tttaatttac gtggtgttga aggatagaa tttetetacac chacaaatag etttacgtat 1200 cqaggaagag gtacagttga tetettaace gaattacege ctgaggataa tagtgtgeta 1260 cctcgcgagg gatatagtca tegtttatgt catgcacet tegttcaaag atetggaaca 1210 ccatttttaa caacagggtg agtatttett tggacgcate gaatgcace tettacaaat 1380 acaattgate cagacaaaat tactcagata cetttagtga aaggattga ggttggggt 1440 ggcgcctctg tegttacagg accaggttt acaggaaggg atatectteg aggaacaaca 1500 tttggggatt tggatcagg accaggttt acaggaaggg atatectteg aggacacaac 1500 tttggggatt tgtatcat gcaagttaat ataaattcac caatacacaa aggatacog 1560 ttaagaattt gttatgette cagtaggat gcacgactta cagtagcgac aggacgaca 1620 aacacaggag ttggagggaa aacattaga tataccgat ttagtaacca tatggaaatt 1680 ggaggaagact taacatctag acatttaga tataccgatt ttagtaacca tatggaaatt 1740 agagtatatc cagatataat tcgtataaat gaacaacaca tattcggtge agacttatt 1800 agatagtgtga aactttata gacaagaag gcggtgaatg cgcgtttac ttetacaaac caaagaggat taaaaacaga tgggacggat tacatattg atacaggat caatttagta 1920 cataggaaga gactagat ttgtgtaga gaaaacagaa gcggtgatag cgcgtttac tettacaaac caaagaggat taaaaacaga tgggacggat tacatattg acaagtatc caatttagta 1920 cataggaaga gactcagtg tgagaggaa attaccacaa aggaggagat 2160 gagagattata cggatgaat ttgtgtaga gaaaacagaa gaggtgattat caaaagaca 1920 cataggaaga gactcagt tgagaggaa tatacatattg acaagtatc caatttagt 1920 cataggaaga gaactagaa tgagaggaa attacacaat 2200 cataggaaga gaacagaa tgagaggaa attacacaa taggagaat 2200 cataggaaga gaacagaa tgagaggaa attacacaa taggagaat 2200 cataggaaga gaacagaa tgagaggaa tacaacacaa agagagaat 2200 cataggaaga gaacagaa tgagaggaa accaagaaa accattagaa 2200 cataggaagg gaacagaa acgaggaa accaagaa accattagaa 2200 cataggaaga gaacagaa accaggaa accaagaa accattagaa 2200 gaaaagatga gaacagaa accaagaa ccaagagaa accaacacaca 2200 gaaaagatga gaacagaa accaagaa ccaagagaa accaacacacac	tattggggag	gacatcgagt	aacttctaac	tatataggag	gaggcaacat	aacatctcct	1020
tttaatttaa gtggtgttga aggatagaa ttttctacac cacaaatag ctttacgtat 1260 cgaggaagag gtacagttga tcttttaacc gaattaccgc ctgaggataa tagtgtgcta 1260 cctcgcgagg gatatagtca tcgtttatgt catgcaactt ttgttcaaag atctggaaca 1320 ccatttttaa caacgggtgt agtattttct tggacgcatc gcagtgcaac tcttacaaat 1380 acaaattgatc cagacaaaat tactcagata cctttagtga aaggatttag agtttggagt 1440 ggcgcctctg tcgttacagg accaggtttt acaggagggg atatccttcg aagaactaac 1560 ttttggggatt ttgtatcatc gcagttaat attaattcac caatcacaca aagataccgt 1560 tttagggatt tggtatgtc cagataaat atcacgat taggaggac aggagagaga 1620 aacacaggag ttggaggga aattagtgtg gacatggctc ttcagaaaac tatggaaatt 1680 ggagagaggct taacatctag acatttaga tataccgat ttaggtacac tgggacact 1740 agagctaatc cagatataat tcgtataaat gaacaaccac tattcggtgc aggctcatt 1740 agagctaatc cagatataat tcgtataaat gaacaaccac tattcggtgc aggctcatt 1890 agatggtggtg aactttaata agataaaatt gaaattatc tagcagatgc acatttgaa 1860 gcagaaatag atttggaagg agacagaaag gcggtgaatg cgctgtttac ttctacaaca 1920 caaagaggat taaaaacaga tgtgacggat tatcatattg atcaagtatc caatttagt 1980 gagtgtttat cggatgaatt ttgtctggat gaaaaacaga aattgtccga gaaagtcaaa 2040 catgcgaage gactcagtga tgagcgaaa ttactcatgg acagtaatt cacatccatt 2100 aatgggcaac tagaccgtgg tggagagga agtacggaat ttaccatcac aggaggaga gagtgtttat caaaacaga tggagcgaaa ttaaccagag atcgaaatt cacatccatt 2100 aatgggcaac tagaccgtg tcgagaagga agtacggaac ttgagagg ctatccaacg gggatatca aagagaatta cgtcacacta ctggggtacct ttgagtgag ctatccaacg gggatatca agagaatga agaccggaa accaggatc ttatagcag agaccgaaa 2220 tatttatata aaaaaatag tgagtcgaaa tcaaggtcc ttatggccgc tttcagtcga aggccaatc gggaaacgta dgaaaccgag tacaggtcc ttatggccgc tttcagtcga aggccaatca 2240 ggaaagaggt gaaaccgaa tcgagaaca ccacacctt gaaggtga tgaaaccaaca 2240 ggaaagaggt tagccagaa accaggaaa tcaggaaa ccaggaaa accattagta 2580 aggaaagagt tagccagaa accaggaaa tcaggaaa ccaggaaaa accattagta 2580 ggaaaacgta gaaacaaaca cacggaaa gaggaaaaa aggggaaacga accaggaaa 2700 ttgcaattcg aaacaaaca cacgagaa ggagaaaaa aggggaaacg aacaataca 2700 ttgcaattag aaacaaaca cacgagaa ggagaaaga agggaaacgg t	atatatggaa	gagaggcgaa	ccaggagcct	ccaagatctt	ttacttttaa	tggacctgtt	1080
cgaggagagg gatacagttga ttotttaacc gaattaccgc ctgaggataa tagtgtgta 1260 octcgcgagg gatatagtca tcgtttatgt catgcaactt ttgttcaaag atctggaaca 1320 occatttttaa caacggggg agtattttct tggacgcatc gcagtgcaac tcttacaaaa 1380 acaattgatc cagacaaaat tactcagata cctttagga aaggatttag agtttggag 1410 ggcgcctctg tcgttacagg accaggttt acaggaggg atatccttcg aagaactaac 1500 tttgggggtt ttgtatctat gcaagttaa attaattcac caataacaca aagataccgt 1560 tttaagaattt tgtatctat gcaagttga gacaggagg atatccttcg aagaactaac 1500 tttaggaggat ttgtatcttc cagtaagaat gcacgactta cagtagcgac aggagcagca 1620 aacacaggag ttggagggca aattagtg gacatggct ttcagtaaca tattgaatat 1680 ggagagaggt taacattaa tcgtataaaat gaacacacac tattcggtgc aggctctatt 1800 agtagtgtg acctttaat acgtaaaat gaacacacac tattcggtgc aggctctatt 1800 agtagtgtgg acctttaata tcgtataaaat gaacaacacac tattcggtgc aggctctatt 1800 agtagtgtgg acctttaata agataaaaat gaaatattg tagcaggag aggctcatt ttgtcatga agaagatatg atttggaagga ggcggagaatga ggcgtgatat ccgaagatga taaaaacaga ttgtgacggat taccatttg atcaagtatc caatttagat 1860 gcagaaatta gatttggaag gacacgaag gggggaaatg cgcgtgttac ttctacaaaa 1920 caaagaggag taaaaacaga ttgtgacggat tatcatattg atcaagtatc caatttagt 1980 gagggtgttata cggatgaat ttgtgctggat gaaaaacaga aattgtccga gaaagtcaaa 2040 catgcgaag gaccaggag cgtgagaaga taacacacac tatcacacac aggaggagat 2160 gacgtattac aagaagatat cggagagaga agaccagaga agtacggaa taccacacac taccacacac taccacacac 2220 tatttatatc aaaaaaataga tgagtcgaaa ttaacacaca aggaggagat 2280 ggggatatac aagagatag aactggaaa accattagaa atctgtcacac aggagagaga accaggaaa accatggaaa accatggaaa tcgaagaccac accacaccttg aatggaaacc taaccacacac 2340 ggaaaagaga ggaagacgaa acaggagaa acatggagaa accaggaaaa tcaaggacac cacacaccttg aatggaaacac taaccacacacacacacacacacacacacacac	tttaggactt	tatcaaatcc	tactttacga	ttattacagc	aaccttggcc	agcaccacca	1140
cetetgegagg gatatagtea tegittatgi catgeacatt tigiteaaag atetggaaca 1320 ceatittiaa caacggggt agtatittet tiggacgeate geagtgeaac tettacaaa 1380 acaattgate cagacaaaat tacteagata cettiagga aaggattag agtitggagt 1440 ggegeetetg tegitacagg accaggitt acaggaggg atateetteg aagaactaac 1500 titiggggatt tigtatetat geaagttaat attaatteae caataacaca aagatacegt 1560 titaaggattt gitatgette cagacaggat geacgactta cagacagaa aggagcagca 1620 aacacaggag tiggagggga aattaggtg gacatggeet ticagaaaac tatggaaatt 1680 ggagaaggag taacactag aacattaga tatacegatt tiagtaatee titticatti 1740 agagactaate cagatataat tegitataaat gaacaacae tatteggig aggeetatt 1800 agatagggg aactitigaaag agcacagaag geggigaatg egettitae tictacaaac 1920 caaagaaggat taaaaacaga tiggacggat tateatatig ateaagatae caattiggat 1980 gagagtittat eggatgaat tiggacggat tateatatig ateaagatae caattiggit 1980 gagtgittat eggatgaat tiggacggat tateatatig ateaagatae caattiggit 1980 gagtgittat eggatgaat tiggacggat tateatatig ateaagatae caattiggit 1980 gagtgittat eggatgaat tiggacggat agaaaacaga aattgecaa gaaagtcaaa 2040 catgegaag gactcagtga tiggacggaa titacecagg ategaaatt cacatecatt 2100 aatgggcaac tagacegtgg etggaggaa gatacggata tiaccateca aggaggagat 2160 gacgtattaa aagagaatta egteacacta etgggtacet tigatgagg etatecaate 2220 tattitate aaaaaataga tiggacgaaa titaaaagcet atacceget atgaataaaa 2280 ggaaagtggg gagaacegaa tegatgcaac ecacacettig aatggaatee taatetagat 2460 tigticetigoa gagaacggaa acatgggaaa ecatagtee etaatetaga 2280 ggaaagtgg gagaacggaa acatggaaa ecatagtee etaatetee etaatetaga 2460 tigticetigoa gagaacggaa acatggaaaa ecagagtaa ecatagatae 2460 tigticetigoa gagaacggaa acaaggagg gagaaaaaaaa ggagagaaa accataagta 2640 ggaaagtgg gagaacgaa acaaggaga acaaggagg gagaaaaaaaa	tttaatttac	gtggtgttga	aggagtagaa	ttttctacac	ctacaaatag	ctttacgtat	1200
ccattittaa caacgggtgt agtattite tggacgcate geagtgeaac tettacaaat 1380 acaattgate cagacaaaat tacteagata cettiagtga aaggattag agtitggagt 1440 ggggcetetg tegttacagg accaggttt acagggagg atateetteg aagaactaac 1500 titiggggatt titgtatetat geaagttaat attaatteae caataacaca aagatacegt 1560 titaaggattt ggtatgette cagtaggagt geacgaetta cagtaggac aggagcagca 1620 aacacaggag titggagggca aattagtgtg gacatggete ticagaaaac tatggaaatt 1680 gggagaagget taacactag aacattiaga tatacegatt tagtaatee titticattt 1740 agagetaate cagatataat tegtataaat gaacaacac tatteggtge aggetetatt 1800 agtagtggtg aacttitata agataaaatt gaaattatte tagcagtge aacattigaa 1860 gcagaatata attiggaaag agcacagaag geggtgaatg egetgtitae titetacaaac 1920 caaagaggat taaaaacaga tigtgacggat tateaattig atcaagtate caattiagtt 1980 agtgttitat eggatgaat titgetggat gaaaaacaga attigcega gaaagteaaa 2040 catgegaag gactcagtga tgagegaaa titactecag ategaaatti cacatecatt 2100 aatgggcac tagacgtgg etggaagga agtacggata titacateca aggaggagat 2160 gacgtattea aagaaataa egteacaca etgggtaeet tigatgagg etatecaaca 2220 tattitatate aaaaaataga tgagtegaaa tataacggata titaccateca aggaggagat 2220 tattitatate aagaaataga tagategaaa atetattaa ticgetacaa tgaatacaaca 2340 ggaaagtgtg gagaacegaa tegatgcaa accacectig aatggaatee taatecaaca 2340 ggaaagtgtg gagaacegaa acaggtee tataggege titeagtega aagtecaate 2400 ggaaagtgtg gagaacegaa acaggtee tataggege titeagtega aagtecaate 2400 ggaaagtgtg gagaacegaa acaggtee taaggtee tagaggaaa ecatagga 2280 gggtatateg aggacggaa acaaggaga actaggtee tegaggagaa accattagaa 2580 acgeaagtgg tagteeggg aacaggaga actaggtee tegaggagaa accattagaa 2580 acgeaagtgg tagteegag acaaggag gagaaaaaaaa aggagagaaa accattagaa 2580 acgeaagagg tagteegag acaagagg gagaaaaaaaa aggagaaaa accattagta 2640 ggagaaagtg tagteegag acaagagg gagaaaaaaaa aggagaaaaaaaaaa	cgaggaagag	gtacagttga	ttctttaacc	gaattaccgc	ctgaggataa	tagtgtgcta	1260
acaattgate cagacaaaat tactcagata cetttagtga aaggatttag agtttggagt 1440 ggggcctctg tcgttacagg accaggttt acaggaggg atatccttcg aagaactaac 1500 tttggggatt ttgtatctat gcaagttaat attaattca caataacaca aagataccgt 1560 ttaagatttc gttatgettc cagtaggag gcacgactta cagtaggga aggaggagca 1620 aacaaggag ttggagggga aattagtgtg gacatggctc ttcagaaaac tatggaaatt 1680 ggagagaggt taacatctag aacatttaga tataccgatt ttagtaatcc tttttcattt 1740 agagctaatc cagatataat tcgtataaat gaacaaccac tattcggtgc aggctctatt 1800 agtagtggtg aactttatat agataaaatt gaaattattc tagcagatgc acatttgaa 1860 gcagaatatg atttggaaag agcacagaag gcggtgaatg cgctgtttac ttctacaaac 1920 caaagaggat taaaaacaga tgtgacggat tatcatattg atcaagtac caatttagtt 1980 gagtgtttat cggatgaatt ttgtctggat gaaaaacaga aattgtccga gaaagtcaaa 2040 catgcgaagc gactcagtga tgagcgaaat ttactcaagg atcgaaatt cacatccatt 2100 aatggggaac tagaccgtgg ctggagaga agtacggata taccatcca aggaggagat 2160 gacgtattca aagagaatta cgtcacacta ctgggtacct ttgatgagtg ctatccaacg 2220 tatttatatc aaaaaaataga tggtcgaaa ttaaaagcct ataccccata aggaggagat 2280 gggatattog aggatagtca agacttagaa atctattaa ttcgctacaa tgcaaaacac 2340 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatc taaccagat 2460 gggaaagtgtg gagaaccgaa tcgatgcaca cacaccttg aatggaatc catactagat 2460 gggaaagtgg gagaacggaa acaaggaaa ctaaggtata tgggggatat caaggttaga ggaaaggagga 2520 agttgttggat gtacagaag acaaggaaa ctaaggtat tgggggatatt caagattaaga 2580 acgcaagatg tagccgaag acaaggaa ctaaggata tcagagttt tcgaaggaga accattagta 2640 ggaaagatg tagcacgaa acaaggaaa ctaaggtat tgggggatatt caagattaaga 2580 acgcaagaag ttagctcgtg gaaaagaag gagaaaaaaa ggaggagaaa accattagta 2640 ggaaaaccgt tagctcgtg gaaaaagag gagaaaaaaa ggaggagaa accattagta 2640 ggaaaaccg tagctcgtg gaaaaagag gagaaaaaaa ggaggaaaa accattagta 2640 ggaaaaccg tagctcgtg gaaaaagag gagaaaaaaa ggaggagaa accattagta 2640 ggaaaaccg tagctcgtg gaaaaagag gagaaaaaa aggagaaa accattagta 2640 ggaaaaccg tagctcgtg gaaaaagag gagaaaaaaa ggaggaaaa accattagta 2700 ttgcaattag aaacaaata cgttacaaa gaggcaaaaa aggagcagaa accattataa 2260 gataaaccg t	cctcgcgagg	gatatagtca	tcgtttatgt	catgcaactt	ttgttcaaag	atctggaaca	1320
gegecetetg tegttacagg accaggttti acaggagggg atatectteg aagaactaace 1500 titiggggatt titgatetat geaagttaat attaatteac caataacaca aagatacecgt 1560 titaagatite gitatgette cagtaggat geacgactia cagtagegac aggagcagca 1620 aacacacaggag titiggagggca aattagtgt gacatggete titaagaaac tatggaaatt 1680 ggagagagget taacatetag aacattaga tataccgatt titigaace tititeatit 1740 agagetaate cagatataat tegtataaat gaacaacaca tatteggie aggetetati 1800 agtagtggtg aacattata agataaaatt gaaattatie tagcagatge acattigaa 1860 gcagaatatag attiggaagg agacaggag gegggaatg cgetgtitae titeacaace 1920 caaagaggat taaaaacaga tiggacggat tateatatig ateaagtate caattigat 1980 gagtgittat eggatgaatt tigtetggat gaaaaacaga geggtgaatg acattigat 2040 catgggaage gactcagtga tiggagggaa attactatag attaccate 2100 aatgggaag gactcagtga tiggaggaaa titaccaagg attaccate 2220 tattitatate aaaaaataga tiggaggaaa attaacaggat titaccateca aggaggagat 2160 gacgtattea aagagaata cgtcacacta ctgggtacet titgatgagtg ctatecaacg 2220 tattitatate aaaaaataga tigagtegaaa titaaaageet ataccegeta tigaattaaga 2280 gggaaacagaa atggecagaa tegagtegaaa tetaagaget titacgetacaa tigaaacaca 2340 ggaaaagtga gagaacaggaa acataggaa accacettg aatggaatee taatetagat 2460 tigticetigaa gagacaggaa acataggaca cacacecttg aatggaatee taatetagat 2460 tigticetigaa gagacaggaa acataggaa acataggaa cacacettg aatggaatee taatetagat 2580 aacgcaagatg ticacagaaca acataggaaa ctaggagtate ticagaagaaa acatataga 2580 aacgcaagatg ticactgcaa acatagaaat ctaggagtic ticagaagaaa acatataga 2580 aacgcaagatg ticactgcaga acataggaa ctaggaaaaaaa gagagaaaaa acatataga 2580 gaaaaacgg ticactgcaga acatagaaat ctaggagtic ticaagaagaa accatagaa 2700 titgcaatataa aaacaaaata cgtitacaaa gagagaaaaaa aggagaaaaa acggaaaaaa 2700 titgcaatataa aaacaaaata cgtitacaaa gagagaaaaaa aggagaaaaa acggaaaaaa 2700 titgcaatacaa aacaaaaata cgtitacaaa gagagaaaaa aggagaaaaaa acggaaaaaa 2700 gaaaaacggg ticataagaa acaatagaa 2700 titgcaatagaa attacaagaa 2700 titgcaatagaa aacaaaaaa 2700 titgcaatagaa aacaaaaaa 2700 titgaaaaaggg ticaatacaa aaacaaaaa 2700 gaaaaagaga aacaaaaaaa 2700 t	ccatttttaa	caacgggtgt	agtattttct	tggacgcatc	gcagtgcaac	tcttacaaat	1380
tttggggatt ttgtatctat gcaagttaat attaattcac caataacaca aagataccgt 1560 ttaagatttc gttatgcttc cagtagagat gcacgactta cagtagcgac aggagcagca 1620 aacacaggag ttggagggca aattagtgtg gacatggctc ttcagaaaac tatggaaatt 1680 ggagagaggct taacatctag aacatttaga tataccgatt ttagtaatcc tttttcattt 1740 aggagtaatc cagatataat tcgtataaat gaacaaccac tattcggtgc aggctctatt 1800 agtagtggtg aactttatat agataaaatt gaaattattc tagcagatgc acatttgaa 1860 gcagaatatg atttggaaga agcacagaag gcggtgaatg cgctgtttac ttctacaaac 1920 caaagaggat taaaaacaga tgtgacggat tatcatattg atcaagtac caatttagtt 1980 gagtgtttat cggatgaatt ttgtctggat gaaaaacaga aattgtccga gaaagtcaaa 2040 catgcgaagc gactcagtga tgagcgaaat ttactcatgg atcacacta aggaggggag gacgtattac aggacggg ctggaggga agtacggata ttacaacca aggaggggat gacgtattca aagaacata cgtcacacta ctgggtacct ttgatgagtg ctatccaacg 2220 taattatatc aaaaaataga tgagtcgaaa ttaaaaagcct ataccegcta tgaattaaga 2280 gggtatatcg aggatagtca agacttagaa atcatttaa ttcgctacaa tgcaaaacac 2340 gaaacagtaa atgtgccagg tacaggtcc ttatggccgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa acaatgtgca catcattctc atcatttcc cttggacatt 2520 gatgttggat gtacagactt aaatgagac caacaccttg agggtgatat caagattaaga 2580 acgcaagatg gtcatgcaag actaggaac ctaggttc tcgaggtga tgctttatt 2520 gatgttggat gtacagact aaatgagaa ctaggata taggtttc tcgaagagaa accattagta 2640 ggagaagcgt tagctcgtg gaaaagagc gagaaaaaa gaggtggga tgctttatt 2760 gtaaactcc aatatgatag attacaag gaggcaaaaa aggctggga tgctttattt 2760 gtaaactcc aatatgatag attacaagt gagacaaaag aagctggga tgctttattt 2760 gtaaactcc aatatgatag attacaagt gagacaaaa aggctgtgga tgctttattt 2760 gtaaactcc aatatgatag attacaagt ggaacaaaa aggctgtgga tgctttatt 2760 gtaaactcc aatatgatag attacaagt ggaacaaaa aggctgtgga tgctttatt 2760 gtaaactcc aatatgatag attacaagt ggagacaaa acgctgtag agacaaaaa 2700 aaagggcatg tcaagaac aaacgacat cggagacg tactccaa ggaacgaga 3000 aaagggcatg tagaagaaca aaacgacat cgttcggtcc ttgttccc ggaacggga 3000 aaagggcatg tagaagaac aaacgacat cgttcggtcc ttgt	acaattgatc	cagacaaaat	tactcagata	cctttagtga	aaggatttag	agtttggagt	1440
ttaagattte gttatgette cagtagagat geacgactta cagtagegac aggageagea 1620 aacacaggag ttggagggca aattagtgtg gacatggete tteagaaaac tatggaaatt 1680 ggagagaget taacatctag aacatttaga tatacegatt ttagtaatee tttteteattt 1740 agagetaate cagatataat tegtataaat gaacaacace tatteggte aggetetatt 1800 agtagtggtg aactttata agataaaatt gaaattatte tageagatge aacatttgaa 1860 gcagaatatg atttggaaag agcacagaag geggtgaatg egetgtttae ttetacaaac 1920 caaagaggat taaaaacaga tgtgacggat tatcatattg atcaagtac caatttagt 1980 gagtgtttat eggatgaatt ttgtetggat gaaaaacaga attgteega gaaagteaaa 2040 catgegaage gacteagtga tgagegaaat ttactecagg ategaaatt cacatcatt 2100 aatgggaage gacteagtga tgagegaaat ttactecagg ategaaatt cacatcatt 2100 aatgggaac tagacegtgg etggagaga agtacggata ttaceateca aggaggagat 2160 gacgtattea aagagaatta egtecacaca etgggtacet ttgatgagtg etatecaacg 2220 tatttatate aaaaaataga tgagtegaaa ttaaaageet ataceegeta tgaattaaga 2280 gggtatateg aggatagtea agacttagaa atetattaa ttegetacaa tgeaaaacac 2340 gaaaacagtaa atgtgecagg tacaggtee ttatggeege ttteagtega aagtecaate 2400 ggaaagtgtg gagaacegaa tegatgeaa ecacacettg aatggaatee taatetagat 2460 tgtteetgea gagaegggaa acataggaa etaagtee ateattee ateattee etagtegat gtacgaagtg 2580 acgeaagatg gteatgeaag actaggaaa etagggtat gggtgatatt caagattaag 2580 acgeaagatg gteatgeaag actaggaaa etagggtat gggggaaaacgg accaaaaca 2700 ttgeaattag aaacaaatat egtttacaaa gaggeaaaag aggtgtgga tgetttattt 2760 gtaaacetee aatatgatag attacaagtg gaaaaaaaa tggagagaaa accattagta 2640 ggagaagegt tageteggag actaggaaa etaggaaa tetagggtte tegaagaga 2820 gataaactee aatatgatag attacaagtg gaaaaaaaa tggagagaaa accattagta 2640 ggagaagegt tecatgeag actaagaag gagaaaaaa aggegtgga tgetttattt 2760 gtaaacetee aatatgaaga attacaagtg gaaaaaaaa ggaggagaa accattagta 2640 ggagaagegt tecatgaaa eeggagaa attacaaa gaggeaaaa acgagaaaa 2700 ttggaaatggg tteatgaaa acaagage tteatecaag ggtateetg gatteeggg 2820 gataaacagag tteatgaaa aaacagaget ttettecag agttateetg gatteeggga 2820 gataacaagag tteatgaaa aaacagacat egtteggtee ttgtteee ggaatggaga 3000 aaagggatgt caca	ggcgcctctg	tegttacagg	accaggtttt	acaggagggg	atatccttcg	aagaactaac	1500
acacacaggag ttggagggca aattagtgg gacatggctc ttcagaaaac tatggaaatt 1680 ggagagggct taacatctag acatttaga tataccgatt ttagtaatcc tttttcattt 1740 agagctaatc cagatataat tcgtataaat gaacaaccac tattcggtgc aggctctatt 1800 agtagtggtg aactttatat agataaaatt gaaattattc tagcagatgc aacatttgaa 1860 gcagaatatg atttggaaag agcacagaag gcggtgaatg cgctgtttac ttctacaaac 1920 caaagaggat taaaaacaga tgtgacggat tatcatattg atcaagtatc caatttagtt 1980 gagtgtttat cggatgaatt ttgtctggat gaaaacagaa aattgtccga gaaagtcaaa 2040 catgcgaagc gactcagtga tgagcgaaat ttactcaagg atcgaaattt cacatccatt 2100 aatgggaac tagaccgtgg ctggagagga agtacggata ttaccatcca aggaggaga 2220 tatttatatc aaaaaaataga tgagtcgaaa ttaaaagcct atacccacca aggaggagag 2220 tatttatatc aaaaaaataga tgagtcgaaa ttaaaagcct ataccgcta tgaattaaga 2280 gggatattca aggatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacac 2340 gaaaacagtaa atgtgccagg tacaggtcc ttatggcgg tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaa ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa acaatggca catcattctc atcattctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggtgtat ggggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctaggtgtat gggggaaacgaa accattagat 2460 tgttcctgca gagacggga acaaggagg gaaaaagaa acaattagat 2640 ggagaagggt tagctgcag actaggaaat ctaggtgttc tcgaagagaa accattagat 2640 ggagaagcgt tagctcgtg gaaaagagg gagaaaaaa aggcgaaaaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaa aggcgagaaa accattagta 2640 ggaaaaccga ttcatgaag actaggaaa ctaggaaat ttcacaggg gagaaaaaa aggcgaaaaa 2700 ttgcaattag aaacaaatat cggttacaaa gaggcaaaaa aggcgaaaaa 2700 ttgcaattag aaacaaata ccggaaagag gtaccaaacac 2820 gataaaacgcg ttcatagaa ccaggaaggg ttacttccag aggttattt tcactgcatt cccctatat 2940 gatacaaacgg ttcatagaa aaacagaccat cgttcggtc ttgttgtccc ggaacgga 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtc ttgttgtccc ggaacgga 3120 gaacaaaag agggatatgg agaaggcatg agaagacga aaacgacat cgttcggtc ttgttgtccc ggaacgaga 3120 gcatacaaaa aggggatatg agaaggctg tacacaaca aaagggcatg agaacaaaca 2320 aaacaacac 3320 a	tttggggatt	ttgtatctat	gcaagttaat	attaattcac	caataacaca	aagataccgt	1560
ggagagagct taacatctag acatttaga tataccgatt ttagtaatcc tttttcattt 1740 agagctaatc cagatataat tcgtataaat gaacaaccac tattcggtgc aggctctatt 1800 agtagtggtg aactttatat agataaaatt gaaattattc tagcagatgc aacatttgaa 1860 gcagaatatg atttggaaag agcacagaag gcggtgaatg cgctgtttac ttctacaaac 1920 caaagaggat taaaaacaga tgtgacggat tatcatattg atcaagtatc caatttagtt 1980 gagtgtttat cggatgaatt ttgtctggat gaaaacagaa aattgtccga gaaagtcaaa 2040 catgcgaagc gactcagtga tgagggaaat ttactcaagg atcgaaatt cacatccatt 2100 aatgggaagc gactcagtga tgagggaaat ttactcaagg atcgaaatt cacatccatt 2100 aatgggaac tagaccgtgg ctgggagaga agtacggata ttaccatcca aggaggagat 22160 gacgtattca aagagaatta cgtcacacta ctgggtacct ttgatgagtg ctatccaacg 2220 tatttatatc aaaaaaataga tgagtcgaaa ttaaaagcct ataccgcta tgaattaaga 2280 gggtatatcg aggatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacac 2340 gaaaacagtaa atgtgccagg tacaggtcc ttatggcgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaac cacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa acaatggca catcattctc atcatttcc cttggacatt 2520 gatgttggat gtacagactt aaatggaagt ctaggtgtat ggggaaagtgt gaaagacgg actaggagaa ctagggaaa ctaggttga ggagaaaggg taacggaaa ctagggaaa ctagggtttc tcgaagagaa accattagat 2640 ggaaagaggt tagctgcag actaggaaat ctaggtttc tcgaagagaa accattagta 2640 ggaaaacggt tagctgcag actaggaaat ctagggtttc tcgaagagaa accattagta 2640 ggaaaacggt tagctcgtg gaaaagagg gaaaaaaaa ggagaaaaaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaag aagctgtgga tgctttattt 2760 gtaaaactctc aatatgatag attacaaagg gatacgaaca ttgccatgat tcatgcggca 2820 gataaaacgcg ttcatagaat ccgagaaggg ttacttccag agttactgt gattccggg 2880 gtcaatagcgg ctattttcga agaattagaa ggggtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatggaa 3000 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttcaa 3120 gcatacaaaa aggggatatgg agaaggctg gaaaggactg ataccaaca 2320 agaagaacaa 2320 agaagaacaa 2320 agaagaacaa 2320 agaagaacaa 23	ttaagatttc	gttatgcttc	cagtagagat	gcacgactta	cagtagcgac	aggagcagca	1620
agagctaatc cagatataat tegtataaat gaacaaccac tatteggtge aggetetatt 1800 agtagtggtg aactttatat agataaaatt gaaattatte tageagatge aacatttgaa 1860 geagaatatg atttggaaag ageacagaag geggtgaatg egetgtttae teetacaaac 1920 caaaagaggt taaaaacaga tgtgaeggat tateatattg ateaagtate caatttagtt 1980 gagtgtttat eggatgaatt ttgteetggat gaaaaacaga aattgeega gaaagteaaa 2040 catgegaage gacteagtga tgagegaaat ttacteeagg ategaaattt cacatecatt 2100 aattgggaac tagacegtgg etggagaga agtacggat ttateataca aggaggagat 2160 gaegtattea aagacegtgg etggagaga attaacagga attgaaggt etateeaacg 2220 tatttatate aaaaaataga tgagegaaa ttaaaagcet ataceegeta tgaattaaga 2280 gggtatatee aggatagtea agacttagaa atetattaa ttegetacaa tgeaaaacac 2340 gaaacagtaa atggeeaga tegatgeea accacettg aatggeeaga aggetaatee 2400 ggaaagtgtg gagaacegaa tegatgeaa eccacacettg aatggaatee taateetagat 2460 tgtteetgea gagaceggaa aacatggea catcatteet ateatteet ettggacatt 2520 gatgtteggat gtacaggete aaagtgea catcatteet ateatteete ettggacatt 2520 gatgtteggat gtacagacat aaatggaaat etaggtgta gggtgatatt eaagattaag 2580 acgeaagatg gteatgeaga actaggaat etaggtgta gggtgatatt eaagattaag 2640 ggaagaageg tageetggt gaaaagagg gagaaaaaaa aggggaaaaa accattagta 2640 ggaagaageg tageetggt gaaaagageg gagaaaaaaa aggggaaaaa accattagta 2640 ggaagaageg tageetggt gaaaagageg gagaaaaaaa aggggaaaaaa accattagta 2640 ggaagaageg tageetggt gaaaagageg gagaaaaaaa aggggaaaaaa accattagta 2640 ggaagaageg teatteegg gaaaagagg gagaaaaaaa aggggaaaaaa 2700 ttgeaaattag aaacaaatat egtttacaaa gaggeaaaaa aagetgtgga tgetttattt 2760 gtaaacetee aatatgatag attacaagtg gatacgaaca ttgeeatgat teateegge 2820 gataaacegeg tteatagaa accatagaa agggtatt teateegge 2820 gataaacegeg tteatagaa aaaagagcat etttaaaa aaatggegat tttaaaaaa gggttattt teacetgeatt etceetata 2940 gatgegagaa atgetataa aaaatggegat tttaaaaaag gggttattt teacetgeat etceetata 2940 gatgegagaa atgetataa aaaatggegat tttaaaaaag gggttateet etggtteee ggaaagaga 3000 aaaagggeat tagaagaaca aaacgaccat egtteggtee ttgttgteee ggaaagaaga 3180 gacaaacaa agggaatatgg agaagagtgt gaaacaacaa accggtaaca	aacacaggag	ttggagggca	aattagtgtg	gacatggctc	ttcagaaaac	tatggaaatt	1680
agtagtggtg aactttatat agataaaatt gaaattatte tagcagatge aacatttgaa 1860 geagaaatatg atttggaaag ageacagaag geggtgaatg egetgttac ttetacaaac 1920 caaaagaggat taaaaacaga tgtgacggat tateatattg ateaagate caatttagtt 1980 gagtgtttat eggatgaatt ttgetetggat gaaaaacaga aattgeega gaaagteaaa 2040 catgegaage gacteagtga tgagegaaat ttacteeagg ategaaattt cacatecatt 2100 aatgggaage gacteagtga tgagegaaga agtacggata ttaccateca aggaggagat 2160 gacgtattea aagagaatta egtecacacta etgggtacet ttgatgagtg etateeaaca 2220 tatttatate aaaaaaataga tgagetegaaa ttaaaaageet ategacgeta tgaaataaga 2280 gggtatateg aggatagtea agacettagaa atetattaa ttegetacaa tgeaaaacac 2340 gaaaacagtaa atgtgeeagg tacaggttee ttatggeege ttteagtega aagteeaate 2460 tgtteetga gagaacegaa tegatgeaa eacacettg aatggaatee taatetagat 2460 tgtteetga gagaacegaa acaatggea eacacettg aatggaatee taatetagat 2460 tgtteetga gagaacegaa acaatggea eacaggttee teaggtgata eacagttaga 2580 acgeaagatg geaaagagg acaagagg gagaaaaaa etgaggtate tegaggagaa accattagta 2640 ggagaagagt tagetegga actaggaaa etaggagat etagggttee teagaggtate tegagagaaa accattagat 2640 ggagaagagg teageeggaaaa actaggaaa etaggagaa accattagaa 2700 ttgeaactag geaaaagag geaaaagag gagaaaaaa aggegaaaaa 2700 ttgeaactee aatatgata 2640 ggagaaacete aaacaaatat egtttacaaa gaggeaaaaa aggeggaaaa accattagta 2640 ggagaaacete aaacaaatat egtttacaaa gaggeaaaaa aggeggaaaa 2700 ttgeaactee aatatgata 2640 agaaagage teagaaaaa eeggaaaaaa 2700 gaaaacege tteatagaat eegagaaag tateteeaa aggegaaaa atgeeggaa 2820 gataaaceete aatatgata 2640 agaacgage tatetteea agaatagaa gggtgattt teacetgat teaceggea 2820 gataaaceete aatatgata aaaagagegg tatetteeaa aggeggaaaa atgeeggaa 3000 aaaagggeatg tagaaaaaa aaacgaceat egteeggee ttgttgeee ggaatggaa 3000 aaaagggeatg tagaaagaaca aaacgaceat egteeggtee ttgttgeee ggaatggaa 3000 aaaagggeatg tagaaagae aaaagaceat egteeggee etatteee tegtgttaca 3120 geatacaaaag agggatatag agaaggetg gaaagagtet atecaaaaaa aagggaagagaagagaagaagaagaagagaaga	ggagagagct	taacatctag	aacatttaga	tataccgatt	ttagtaatcc	tttttcattt	1740
gcagaatatg atttggaaag agcacagaag gcggtgaatg cgctgtttac ttctacaaac 1920 caaaagaggat taaaaacaga tgtgacggat tatcatattg atcaagtatc caatttagtt 1980 gagtgtttat cggatgaatt ttgtctggat gaaaaacagag aattgtccga gaaagtcaaaa 2040 catgcgaage gactcagtga tgagcgaaat ttactccagg atcgaaattt cacatccatt 2100 aatgggaac tagaccgtgg ctggagagga agtaccggata ttaccatcca aggaggagat 2160 gacgtattca aagagaatta cgtcacacta ctgggtacct ttgatgagtg ctatccaacg 2220 tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct atacccgcta tgaattaaga 2340 gaaacagtaa atgtgccagg tacaggttcc ttatggcgc tttcagtcga aagtccaatc 2400 ggaaacggtg gagaaccgaa tcgatgcaa ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagaacggaa accatgtgca catcattct atcattctc cttggacatt 2520 gatgttggat gtacaggct aacaggtcc catcattctc atcattctc cttggacatt 2520 gatgttggat gtacagact aacaggaaa ctaggaaac ctagggttc tcgagagagaa accattagta 2640 ggagaagggt tagccagga accaggaaa ctaggaaac ctagggttc tcgagagagaa accattagta 2640 ggagaagggt tagctcgtgt gaaaagaggg gagaaaaaat ggagagacaa accattagta 2640 ggagaagggt tagctcgtgt gaaaagagg gagaaaaaa ctagggaaaa accattagta 2640 ggagaagggt tagctcgtgt gaaaagagg gagaaaaaa ttggagagaa accattagta 2640 gtaaactcc aatagtaga attacaagt gagaaaaaa aggggaaaaa accattagta 2640 gtaaactcc aatagtaga attacaagt gagaaaaaa ttgccatga tcatgcgga 2820 gataaaccgg ttcatagaat ccggaaaggg tatcttccag agttatctgt gattccgggc 2820 gataaacccc aatattca aatagaag attacaagt gaaacaaaa acgcgaaaaa 2700 ttgcaattag aacaaaatat cgtttacaaa gaggcaaaaa ttgccatgat tcatgcggca 2820 gataaaccgg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgcattaa aaatggcgat tttaataatg gcttatcatg ctggaacggg 3000 aaaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3000 aaaagggcatg tagaagaac aaacgaccat cgttcggtcc ttgttgtccc ggaatggaa 3180 gcatacaaaag agggatatgg agaaggctg gtaaccattc atgggatcaa accaggtaaca 3180 gacaaactga agtttagaa ctgcgtagaa aggagatgt agcaaacaa accaggtaacgaa aggacaacaa agggataaa accaggaaagagaagaagaagaagaagaagaagaagaaga	agagctaatc	cagatataat	tcgtataaat	gaacaaccac	tattcggtgc	aggctctatt	1800
caaagaggat taaaaacaga tgtgacggat tatcatattg atcaagtatc caatttagtt 1980 gagtgtttat cggatgaatt ttgtctggat gaaaacagag aattgtccga gaaagtcaaaa 2040 catgcgaage gactcagtga tgagcgaaat ttactccagg atcgaaattt cacatccatt 2100 aatgggaac tagaccgtgg ctggagagga agtacggata ttaccatcca aggaggagat 2160 gacgtattca aagacgtgg ctggagagga agtacggata ttaccatcca aggaggagat 2160 gacgtattca aagaagaatta cgtcacacta ctgggtacct ttgatgagtg ctatccaacg 2220 tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct atacccgcta tgaattaaga 2280 gggtatatcg aggatagtca agacttagaa atctattaa ttcgctacaa tgcaaaacac 2340 gaaacagtaa atgtgccagg tacaggttcc ttatggccgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa aacatgtgca catcattctc atcatttctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggtttc tcgaaggagaa accattagta 2640 ggagaagagt gtcatgcaag actaggaaat ctaggtttc tcgaaggagaa accattagta 2640 ggagaagagt gtcatgcaag actaggaaat ctaggtttc tcgaaggagaa accattagta 2640 ggagaagagt tagctcgatg gaaaagagag gagaaaaaa gagggaaaaa accattagta 2640 ggagaagcgt tagctcgatg gaaaagagcg gagaaaaaa gagggaaaaa accattagta 2640 ggagaagcgt tagctcgatg gaaaagagcg gaaaaaaaa gaggaaaaaa accattagta 2640 ggaaaaccg ttcatagaaa ccggaaagcg tatctccaa agaattagaa gaggaaaaaa ttgccatgat tcatgcggca 2820 gataaacccc aatatgaa attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaacccg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgccattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaaa agggatatgg agaaggctg gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagaa ctgcgtagaa agggataga aggtaacaa ccgcgtaacg 3240	agtagtggtg	aactttatat	agataaaatt	gaaattattc	tagcagatgc	aacatttgaa	1860
gagtgtttat cggatgaatt ttgtctggat gaaaaacgag aattgtccga gaaagtcaaa 2040 catgcgaagc gactcagtga tgagcgaaat ttactccagg atcgaaattt cacatccatt 2100 aatgcgaagc gactcagtga tgagcgaaat ttactccagg atcgaaattt cacatccatt 2100 gacgtattca aagacgtgg ctggagagga agtacggata ttaccatcca aggaggagat 2160 gacgtattca aagagaatta cgtcacacta ctgggtacct ttgatgagtg ctatccaacg 2220 tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct atacccgcta tgaattaaga 2280 gggatatatcg aggatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacac 2340 gaaacagtaa atgtgccagg tacaggttcc ttatggccgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa aacatgtgca catcattctc atcattctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggtgtat gggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctaggtttc tcgaagagaa accattagta 2640 ggagagaagcgt tagctcgtgt gaaaagagcg gagaaaaaaa ggaggagaaa accattagta 2640 ggagagagcgt tagctcgtgt gaaaagagcg gagaaaaaaa ggaggagaaa accattagta 2640 ggagaaacgt tagctcgtgt gaaaagagcg gagaaaaaa ggaggagaaa accattagta 2640 ggaaaacgg ttcatgcaag actaggaaat ctagggtttc tcgaagagaaa accggaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaa aggctgtgga tgctttattt 2760 gtaaacctcc aatatgatag attacaaagtg gatacgaaca ttgccatgat tcatgcggc 2820 gataaaccgg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataaa ggcttatcat ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataaa ggcttatcat ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tctaatagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgt gctatatcct tcgtgttaca 3120 gcatacaaaa agggaatatgg agaaggctgt gtaaccatc atgggatca accaggaacaa accggaacaaa accggaacaga aggaacaaaa agggaacaaaa agggaaaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gagaggctgt ataccaatc atgggatcaacaac 3240	gcagaatatg	atttggaaag	agcacagaag	gcggtgaatg	cgctgtttac	ttctacaaac	1920
catgogaago gactcagtga tgagogaaat ttactccagg atcgaaattt cacatccatt 2100 aatgggcaac tagaccgtgg ctggagagga agtacggata ttaccatcca aggaggagat 2160 gacgtattca aagagaatta cgtcacacta ctgggtacct ttgatgagtg ctatccaacg 2220 tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct atacccgcta tgaattaaga 2280 gggtatatcg aggatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacac 2340 gaaacagtaa atgtgccagg tacaggttcc ttatggccgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacggaa aacatgtgca catcattctc atcattctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggtgtat gggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctagagtttc tcgaagagaa accattagta 2640 ggaaaacgtt tagctcgtg gaaaagagcg gagaaaaaa gggagaacaa accattagta 2640 ggaaaacgt tagctcgtg gaaaagagcg gagaaaaaa gggagagacaa accattagta 2640 ggaaaactcc aatatgatag attacaaagt gatacgaaca ttgccatgat tcatgcggca 2820 gataaactcc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaacgg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccatcc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	caaagaggat	taaaaacaga	tgtgacggat	tatcatattg	atcaagtatc	caatttagtt	1980
aatgggcaac tagaccgtgg ctggagagga agtacggata ttaccatcca aggaggagat 2160 gacgtattca aagagaatta cgtcacacta ctgggtacct ttgatgagtg ctatccaacg 2220 tatttatatc aaaaaataga tgagtcgaaa ttaaaaagcct ataccgcta tgaattaaga 2280 gggtatatcg aggatagtca agacttagaa atctattaa ttcgctacaa tgcaaaacac 2340 gaaacagtaa atgtgccagg tacaggtcc ttatggcgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa aacatgtgca catcattctc atcattctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggttat gggtgatatt caaggattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctaggtttc tcgaagagaa accattagta 2640 ggagaaagcgt tagctcgtgt gaaaagagcg gagaaaaaat ggagagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaa aggctgga tgctttattt 2760 gtaaactctc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaacgg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aacgaccat cgttcggtc ttgttgtccc ggaatgggaa 3060 gcagaaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	gagtgtttat	cggatgaatt	ttgtctggat	gaaaaacgag	aattgtccga	gaaagtcaaa	2040
gacgtattca aagagaatta cgtcacacta ctgggtacct ttgatgagtg ctatccaacg 2220 tatttatat aaaaaataga tgagtcgaaa ttaaaagcc ataccgcta tgaattaaga 2280 gggtatatcg aggatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacac 2340 gaaacagtaa atgtgccagg tacaggtcc ttatggccgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa aacatgtgca catcattctc atcatttcc cttggacatt 2520 gatgttggat gtcatgcaag acatggaaat ctaggtgtat gggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctaggtttc tcgaagagaa accattagta 2640 ggagagagcg tagctcgtgt gaaaagagcg gagaaaaaaa ggagagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaa aggctgtga tgctttattt 2760 gtaaactctc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcgga 2820 gataaaccgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataag ggttatctgt gattccggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaaa agggaatatg agaagacgtg gaaacaataca cacggtaacg 3240 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa aacggaacaa agggatatga agaagacgtg gaaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	catgcgaagc	gactcagtga	tgagcgaaat	ttactccagg	atcgaaattt	cacatccatt	2100
tatttatatc aaaaaataga tgagtcgaaa ttaaaagcct atacccgcta tgaattaaga 2280 gggtatatcg aggatagtca agacttagaa atctatttaa ttcgctacaa tgcaaaacac 2340 gaaaacagtaa atgtgccagg tacaggttcc ttatggccgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa aacatgtgca catcattctc atcattctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggtgtat gggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctaggtgtat gggtgatatt caagattaag 2640 ggagaaagcgt tagctcgtgt gaaaaagagcg gagaaaaaaa ggagagacaa accattagta 2640 ggagaaagcgt tagctcgtgt gaaaagagcg gagaaaaaaa ggagagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaa aggctgtgga tgctttattt 2760 gtaaacctcc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaacgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccatc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	aatgggcaac	tagaccgtgg	ctggagagga	agtacggata	ttaccatcca	aggaggagat	2160
gggtatatcg aggatagtca agacttagaa atctattaa ttcgctacaa tgcaaaacac 2340 gaaacagtaa atgtgccagg tacaggttcc ttatggccgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa aacatgtgca catcattctc atcattctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggtgtat gggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctaggtgtat tcgaagagaa accattagta 2640 ggagaagcgt tagctcgtg gaaaagagcg gagaaaaaat ggagagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaa aggcggaaaaa tcattgga 2760 gtaaactctc aatatgatag attacaaag ggagaaacaa ttgccatgat tcatgcggca 2820 gataaaaccgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaagctgt gtaaccattc atgggatcga gaacaataca 3190 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	gacgtattca	aagagaatta	cgtcacacta	ctgggtacct	ttgatgagtg	ctatccaacg	2220
gaaacagtaa atgtgccagg tacaggttcc ttatggccgc tttcagtcga aagtccaatc 2400 ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa aacatgtgca catcattctc atcatttctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggtgtat gggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctaggtttc tcgaagagaa accattagta 2640 ggagaagcgt tagctcgtgt gaaaagagcg gagaaaaaa ggagagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaa aagctgtgga tgctttattt 2760 gtaaactctc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaaccgg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgccattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gacaaccaga agtttagcaa ctgcgtagaa gagaagtct atccaaacaa cacggtaacg 3240 gacagaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	tatttatatc	aaaaaataga	tgagtcgaaa	ttaaaagcct	atacccgcta	tgaattaaga	2280
ggaaagtgtg gagaaccgaa tcgatgcaca ccacaccttg aatggaatcc taatctagat 2460 tgttcctgca gagacgggaa aacatgtgca catcattctc atcattctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggtgtat gggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctagagtttc tcgaagagaa accattagta 2640 ggagaagcgt tagctcgtgt gaaaaagacg gagaaaaaat ggagagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaag aagctgtgga tgctttattt 2760 gtaaactctc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaaccgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	gggtatatcg	aggatagtca	agacttagaa	atctatttaa	ttcgctacaa	tgcaaaacac	2340
tgttcctgca gagacgggaa aacatgtgca catcattctc atcattctc cttggacatt 2520 gatgttggat gtacagactt aaatgaagat ctaggtgtat gggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctaggtttc tcgaaggaaa accattagta 2640 ggagagaagcgt tagctcgtgt gaaaagagcg gagaaaaaat ggagagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaag aagctgtgga tgctttattt 2760 gtaaaactctc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaaacgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	gaaacagtaa	atgtgccagg	tacaggttcc	ttatggccgc	tttcagtcga	aagtccaatc	2400
gatgttggat gtacagactt aaatgaagat ctaggtgtat gggtgatatt caagattaag 2580 acgcaagatg gtcatgcaag actaggaaat ctagagtttc tcgaagagaa accattagta 2640 ggagaaagcgt tagctcgtgt gaaaaagacg gagaaaaaat ggagagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaag aagctgtgga tgctttattt 2760 gtaaactctc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaacgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	ggaaagtgtg	gagaaccgaa	tcgatgcaca	ccacaccttg	aatggaatcc	taatctagat	2460
acgcaagatg gtcatgcaag actaggaaat ctagagtttc tcgaagagaa accattagta 2640 ggagaagcgt tagctcgtgt gaaaagagcg gagaaaaaa ggaggagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaaa aagctgtgga tgctttattt 2760 gtaaactotc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaacgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	tgttcctgca	gagacgggaa	aacatgtgca	catcattctc	atcatttctc	cttggacatt	2520
ggagaaagcgt tagctcgtgt gaaaaagacg gagaaaaaat ggagagacaa acgcgaaaaa 2700 ttgcaattag aaacaaatat cgtttacaaa gaggcaaaag aagctgtgga tgctttattt 2760 gtaaactctc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaacgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	gatgttggat	gtacagactt	aaatgaagat	ctaggtgtat	gggtgatatt	caagattaag	2580
ttgcaattag aaacaatat cgtttacaaa gaggcaaaag aagctgtgga tgctttattt 2760 gtaaactctc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaacgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	acgcaagatg	gtcatgcaag	actaggaaat	ctagagtttc	tcgaagagaa	accattagta	2640
gtaaactctc aatatgatag attacaagtg gatacgaaca ttgccatgat tcatgcggca 2820 gataaacgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	ggagaagcgt	tagctcgtgt	gaaaagagcg	gagaaaaaat	ggagagacaa	acgcgaaaaa	2700
gataaacgcg ttcatagaat ccgagaagcg tatcttccag agttatctgt gattccgggt 2880 gtcaatgcgg ctattttcga agaattagaa gggtgtattt tcactgcatt ctccctatat 2940 gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	ttgcaattag	aaacaaatat	cgtttacaaa	gaggcaaaag	aagctgtgga	tgctttattt	2760
gtcaatgogg ctattttoga agaattagaa gggtgtattt toactgoatt otooctatat 2940 gatgogagaa atgtoattaa aaatggogat tttaataatg gottatoatg otggaacgtg 3000 aaagggoatg tagaagaaca aaacgaccat ogttoggtoo ttgttgtooc ggaatgggaa 3060 goagaagtgt cacaagaggt togtgtotgt ocaggtogtg gotatatoot togtgttaca 3120 goatacaaag agggatatgg agaaggotgt gtaaccatto atgggatoga gaacaataca 3180 gacgaactga agtttagoaa otgogtagaa gaggaagtot atooaaacaa cacggtaacg 3240	gtaaactctc	aatatgatag	attacaagtg	gatacgaaca	ttgccatgat	tcatgcggca	2820
gatgcgagaa atgtcattaa aaatggcgat tttaataatg gcttatcatg ctggaacgtg 3000 aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	gataaacgcg	ttcatagaat	ccgagaagcg	tatcttccag	agttatctgt	gattccgggt	2880
aaagggcatg tagaagaaca aaacgaccat cgttcggtcc ttgttgtccc ggaatgggaa 3060 gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	gtcaatgcgg	ctattttcga	agaattagaa	gggtgtattt	tcactgcatt	ctccctatat	2940
gcagaagtgt cacaagaggt tcgtgtctgt ccaggtcgtg gctatatcct tcgtgttaca 3120 gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	gatgcgagaa	atgtcattaa	aaatggcgat	tttaataatg	gcttatcatg	ctggaacgtg	3000
gcatacaaag agggatatgg agaaggctgt gtaaccattc atgggatcga gaacaataca 3180 gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	aaagggcatg	tagaagaaca	aaacgaccat	cgttcggtcc	ttgttgtccc	ggaatgggaa	3060
gacgaactga agtttagcaa ctgcgtagaa gaggaagtct atccaaacaa cacggtaacg 3240	gcagaagtgt	cacaagaggt	tcgtgtctgt	ccaggtcgtg	gctatatcct	tcgtgttaca	3120
	gcatacaaag	agggatatgg	agaaggctgt	gtaaccattc	atgggatcga	gaacaataca	3180
tgtaatgatt atcctgcaaa tcaagaagaa tacgggggtg cgtacacttc tcgtaatcgt 3300	gacgaactga	agtttagcaa	ctgcgtagaa	gaggaagtct	atccaaacaa	cacggtaacg	3240
	tgtaatgatt	atcctgcaaa	tcaagaagaa	tacgggggtg	cgtacacttc	tcgtaatcgt	3300

-continue	

ggatatgacg aaacttatgg aagcaattet teegtateag etgattatge gteagtttat gaagaaaaag cgtatacaga tggacgaaga gacaatccat gtgaatttaa cagagggtat ggggattata cgccactacc agctggctat gtaacaaaag aattagaata cttcccagaa accgataagg tatggattga gattggagaa acggaaggaa cattcatcgt ggacagtgtg gaattactcc ttatggagga ataa <210> SEQ ID NO 6 <211> LENGTH: 1187 <212> TYPE: PRT <213> ORGANISM: Bacillus thuringienses <220> FEATURE: <221> NAME/KEY: MISC_FEATURE <222> LOCATION: (1)..(1187) <223> OTHER INFORMATION: The amino acid sequence of the TIC1425 pesticidal protein. <400> SEQUENCE: 6 Met Asn Asn Ile Glu Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser Asn Pro Glu Glu Val Ile Leu Asp Gly Glu Arg Ile Ser Thr Gly Asn Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Leu Leu Val Ser Asn 40 Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val Trp Gly Ile Val Gly Pro Ser Pro Trp Asp Ala Phe Leu Val Gln Ile Glu Gln Leu Ile Gln Gln Arg Ile Glu Ala Tyr Ala Arg Ala Ala Ala Ile Ser Asn Leu Glu Gly Ile Gly Asn Asn Phe Asn Ile Tyr Val Glu Ala 105 Phe Gln Glu Trp Glu Glu Asp Pro Asn Asn Pro Ala Thr Arg Asn Arg Val Val Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp Ile 135 Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val Tyr Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val Ile Phe Gly Glu Arg Trp Gly Leu Thr Thr Thr Asn Val Asn Glu Asn Tyr Asn Arg Gln Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala Asn Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln Asp 215 Trp Ile Thr Tyr Asn Arg Leu Arg Arg Glu Leu Thr Leu Thr Val Leu Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro Ile 250 Gln Pro Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu Ile 265 Thr Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe Asn 280 Val Met Glu Ser Asn Ala Ile Arg Asn Pro His Leu Phe Asp Ile Leu

295

-continued

Asn 305	Asn	Leu	Thr	Ile	Phe 310	Thr	Asp	Trp	Phe	Ser 315	Val	Gly	Arg	Asn	Phe 320
Tyr	Trp	Gly	Gly	His 325	Arg	Val	Thr	Ser	Asn 330	Tyr	Ile	Gly	Gly	Gly 335	Asn
Ile	Thr	Ser	Pro 340	Ile	Tyr	Gly	Arg	Glu 345	Ala	Asn	Gln	Glu	Pro 350	Pro	Arg
Ser	Phe	Thr 355	Phe	Asn	Gly	Pro	Val 360	Phe	Arg	Thr	Leu	Ser 365	Asn	Pro	Thr
Leu	Arg 370	Leu	Leu	Gln	Gln	Pro 375	Trp	Pro	Ala	Pro	Pro 380	Phe	Asn	Leu	Arg
Gly 385	Val	Glu	Gly	Val	Glu 390	Phe	Ser	Thr	Pro	Thr 395	Asn	Ser	Phe	Thr	Tyr 400
Arg	Gly	Arg	Gly	Thr 405	Val	Asp	Ser	Leu	Thr 410	Glu	Leu	Pro	Pro	Glu 415	Asp
Asn	Ser	Val	Leu 420	Pro	Arg	Glu	Gly	Tyr 425	Ser	His	Arg	Leu	Cys 430	His	Ala
Thr	Phe	Val 435	Gln	Arg	Ser	Gly	Thr 440	Pro	Phe	Leu	Thr	Thr 445	Gly	Val	Val
Phe	Ser 450	Trp	Thr	His	Arg	Ser 455	Ala	Thr	Leu	Thr	Asn 460	Thr	Ile	Asp	Pro
Asp 465	Lys	Ile	Thr	Gln	Ile 470	Pro	Leu	Val	Lys	Gly 475	Phe	Arg	Val	Trp	Ser 480
Gly	Ala	Ser	Val	Val 485	Thr	Gly	Pro	Gly	Phe 490	Thr	Gly	Gly	Asp	Ile 495	Leu
Arg	Arg	Thr	Asn 500	Phe	Gly	Asp	Phe	Val 505	Ser	Met	Gln	Val	Asn 510	Ile	Asn
Ser	Pro	Ile 515	Thr	Gln	Arg	Tyr	Arg 520	Leu	Arg	Phe	Arg	Tyr 525	Ala	Ser	Ser
Arg	Asp 530	Ala	Arg	Leu	Thr	Val 535	Ala	Thr	Gly	Ala	Ala 540	Asn	Thr	Gly	Val
Gly 545	Gly	Gln	Ile	Ser	Val 550	Asp	Met	Ala	Leu	Gln 555	Lys	Thr	Met	Glu	Ile 560
Gly	Glu	Ser	Leu	Thr 565	Ser	Arg	Thr	Phe	Arg 570	Tyr	Thr	Asp	Phe	Ser 575	Asn
Pro	Phe	Ser	Phe 580	Arg	Ala	Asn	Pro	Asp 585	Ile	Ile	Arg	Ile	Asn 590	Glu	Gln
Pro	Leu	Phe 595	Gly	Ala	Gly	Ser	Ile 600	Ser	Ser	Gly	Glu	Leu 605	Tyr	Ile	Asp
ГÀа	Ile 610	Glu	Ile	Ile	Leu	Ala 615	Asp	Ala	Thr	Phe	Glu 620	Ala	Glu	Tyr	Asp
Leu 625	Glu	Arg	Ala	Gln	630 Lys	Ala	Val	Asn	Ala	Leu 635	Phe	Thr	Ser	Thr	Asn 640
Gln	Arg	Gly	Leu	Lys 645	Thr	Asp	Val	Thr	Asp 650	Tyr	His	Ile	Asp	Gln 655	Val
Ser	Asn	Leu	Val 660	Glu	GÀa	Leu	Ser	Asp 665	Glu	Phe	CÀa	Leu	Asp 670	Glu	Lys
Arg	Glu	Leu 675	Ser	Glu	ГЛа	Val	680 Lys	His	Ala	Lys	Arg	Leu 685	Ser	Asp	Glu
Arg	Asn 690	Leu	Leu	Gln	Asp	Arg 695	Asn	Phe	Thr	Ser	Ile 700	Asn	Gly	Gln	Leu
Asp 705	Arg	Gly	Trp	Arg	Gly 710	Ser	Thr	Asp	Ile	Thr 715	Ile	Gln	Gly	Gly	Asp 720

-continued

_															
Asp	Val	Phe	Lys	Glu 725	Asn	Tyr	Val	Thr	Leu 730	Leu	Gly	Thr	Phe	735	Glu
CÀa	Tyr	Pro	Thr 740	Tyr	Leu	Tyr	Gln	Lys 745	Ile	Asp	Glu	Ser	750		Lys
Ala	Tyr	Thr 755	Arg	Tyr	Glu	Leu	Arg 760	Gly	Tyr	Ile	Glu	Asp 765		Gln	Asp
Leu	Glu 770	Ile	Tyr	Leu	Ile	Arg 775	Tyr	Asn	Ala	Lys	His 780	Glu	Thr	Val	Asn
Val 785	Pro	Gly	Thr	Gly	Ser 790	Leu	Trp	Pro	Leu	Ser 795	Val	Glu	. Ser	Pro	Ile 800
Gly	Lys	Сув	Gly	Glu 805	Pro	Asn	Arg	Cys	Thr 810	Pro	His	Leu	. Glu	Trp 815	Asn
Pro	Asn	Leu	Asp 820	CAa	Ser	Cys	Arg	Asp 825	Gly	Lys	Thr	Cys	Ala 830		His
Ser	His	His 835	Phe	Ser	Leu	Asp	Ile 840	Asp	Val	Gly	Cys	Thr 845		Leu	Asn
Glu	Asp 850	Leu	Gly	Val	Trp	Val 855	Ile	Phe	Lys	Ile	860 Lys		Gln	Asp	Gly
His 865	Ala	Arg	Leu	Gly	Asn 870	Leu	Glu	Phe	Leu	Glu 875	Glu	Lys	Pro	Leu	Val 880
Gly	Glu	Ala	Leu	Ala 885	Arg	Val	Lys	Arg	Ala 890	Glu	Lys	Lys	Trp	Arg 895	Asp
Lys	Arg	Glu	900 Lys	Leu	Gln	Leu	Glu	Thr 905	Asn	Ile	Val	Tyr	Lys 910		Ala
Lys	Glu	Ala 915	Val	Asp	Ala	Leu	Phe 920	Val	Asn	Ser	Gln	Tyr 925		Arg	Leu
Gln	Val 930	Asp	Thr	Asn	Ile	Ala 935	Met	Ile	His	Ala	Ala 940	Asp	Lys	Arg	Val
His 945	Arg	Ile	Arg	Glu	Ala 950	Tyr	Leu	Pro	Glu	Leu 955	Ser	Val	Ile	Pro	Gly 960
Val	Asn	Ala	Ala	Ile 965	Phe	Glu	Glu	Leu	Glu 970	Gly	CAa	Ile	Phe	Thr 975	Ala
Phe	Ser	Leu	Tyr 980	Asp	Ala	Arg	Asn	Val 985		Lys	Asn	Gly	990		Asn
Asn	Gly	Leu 995	Ser	Сув	Trp	Asn	Val		s Gl	y Hi	s Va		u G 05	lu G	ln Asn
	His 1010		g Sei	r Val	l Leu	101		al P	ro G	lu T		lu 020		Glu	Val
Ser	Gln 1025		ı Val	l Arg	y Val	. Cy:		ro G	ly A	rg G		yr 035	Ile	Leu	Arg
Val	Thr 1040		а Туг	. Lys	g Glu	104		yr G	ly G	lu G		ys 050	Val	Thr	Ile
His	Gly 1055		e Glu	ı Asr	n Asr	106		ap G	lu L	eu L		he 065	Ser	Asn	Cys
Val	Glu 1070		ı Glu	ı Val	L Tyr	Pro 10		sn A	sn T	hr V		hr 080	Cys	Asn	Asp
Tyr	Pro 1085		a Asr	n Glr	n Glu	Gli 109		yr G	ly G	ly A		yr 095	Thr	Ser	Arg
Asn	Arg	-	/ Туз	r Asp	Glu	110		yr G	ly S	er A		er 110	Ser	Val	Ser
Ala	Asp 1115	_	Ala	a Sei	. Val	. Ty:		lu G	lu L	Aa Y		yr 125	Thr	Asp	Gly
Arg	Arg	Asp) Asr	n Pro	Cys	; Glı	ı Pl	he A	sn A	rg G	ly T	yr	Gly	Asp	Tyr

		US 11,312,752	2 B 2	
	61			62
		-continued		
1130	1135	1140		
Thr Pro Leu Pro Al	a Gly Tyr Val Thr 1150	Lys Glu Leu Glu Tyr F 1155	Phe	
Pro Glu Thr Asp Ly 1160	rs Val Trp Ile Glu 1165	Ile Gly Glu Thr Glu C	Sly	
Thr Phe Ile Val As	p Ser Val Glu Leu 1180	Leu Leu Met Glu Glu 1185		
	sc_feature (3537) ATION: Nucleic acid	s sequence encoding a TI n Bacillus thuringiensi		
<400> SEQUENCE: 7				
atggataaca atatcaag	gaa ccaatgcatt cctta	acaatt gtttaaatga teete	gaggta 60	
gaaatattag gtgaagag	ggg gataactact agtaa	atgaaa atctcgaatt tttct	tatog 120	
ctaacgaaat ttgtcttg	gaa taggtttgtc cctgg	gtggag catatgtagc tggco	tattt 180	
gatgtattct ggggatgg	itt aaaaccttct gatto	ggtetg caateettga acaaa	attgaa 240	
gaattaatta accaaaaa	at tgagacgttt gctag	gaaatc aagcaattag tagat	tggaa 300	
gggttaagca acctttat	ga aatttacgca gatac	etttta aagaatggga aaaag	gateeg 360	
actaatccag cattaaga	ıga agaaatgegt acaca	aattta atgacatgaa cagct	ctttt 420	
gtaacagcta tgcctctt	tt ttcagttcaa aatti	ttgaag tteetetttt ageag	gtatac 480	
gctcaagctg caaattta	ıca tctatcagtt ttaaç	gggatg teteagtttt tggte	caaaat 540	
tggggatttg attcagco	ac tgtcaatagt cgtta	ataatg atttagtaag aaata	ittegt 600	
acctatacaa attatgto	gt acgttggtat aacad	caggat tagcaaggtt acgag	ggtact 660	
acgtaccaag attggtta	ıaa ttatcatcgc tttaç	gaagag aattaacaat aactg	gcattg 720	
gatatcatta ccatatto	cc acactacgat aataa	aaatgt atccaattca acccc	atttt 780	

960 1020

1080

1140

1200

1320 1380

1440

1500

1560

1620

caattaacaa gagagattta tacggatcca ctaattaatt tcaatccggc gttacagtct

gtagcacaat tacctctatt taatgagatg gaaaatagta caattagaag ccctcattta gttgattttt taaataggct tacaatttat acagattggt atagtctcgg aagacactat

tattggggag gacatcaaat agtctctaga caaacaggat caacttccac tattacattc

cctatatatg gaagagggc gaatcaagag gcccctagaa catataattt tagtcaacct gtctttagaa cactgtcaaa tcctacttta acacgtttaa tgcaaccttg gccagcccca

gcatttcagt tgcgtcgtct tgagggagtt gaatttcaaa caactacagg taattttacg

tatcgaggaa gaggtacggt agattccttt gatgaattac caccagatga tacaagcgta ccagcgcgtg aaggatatag tcatcgttta tgtcatgcaa catttatcag aaaatctggg

acgccgtatt taacaacggg tgtaacacta tcttggacac acaatagcaa tacacctacg
aatataattt atcctgataa aatcactcaa gtaccattgg tgaaagcatc taaccttcat

tctagtgctt tcgttttaaa aggaccagga tttacaggag gggacatact tggaagaact

agtgtgggca acatagcaga tatccaaatg aatattactg caccgttatc acaaagatat

cgcgtacgaa ttcggtatgc ctctactaca aacttacaat ttcatacgac aattaacggc

-continued

```
agggccgtaa atcaggctaa tttcccagca actatgaata gagtagaaga cttagaatat
                                                                    1680
aattoottta gaacgataag tttcggtact cottttaact ttttagatgo tcaaagtaco
                                                                    1740
ttcaggttag gtgtatggag cttttcttca ggtaccgttt taatagatag aattgaagtt
                                                                    1800
gtaccaatgg aagtaacatt tgaagcagaa tctgatttag aaagagcaca aaaggcggtg
                                                                    1860
aacgctctgt ttacttctat aaatcaaaaa ggactaaaaa cagatgtaac agattatcac
                                                                    1920
attgatcaag tatccaattt ggtcgaatgt ctatccgatg aattttgtct agatgaaaag
                                                                    1980
agagaactat ttgagaaagt caaatatgcg aaacgactca gtgacgaacg gaatttactt
                                                                    2040
gcagatccaa atttcacatc tattaatggg caactagatc gtggatggag aggaagtacg
                                                                    2100
gatattacca ttcaaggggg cgatgacgta ttcaaagaaa actacgtcac actatcaggt
accettgatg agtgttatce aacctattta tatcaaaaaa tagacgaate gaaattaaaa
gcgtataccc gttacgagtt acgaggatat atcgaagata gtcaagattt agaagtatac
                                                                    2280
ttgattcqtt acaataccaa acatgaaaca ttgaatgtac caggtacagg gggcctatgg
                                                                    2340
cegettgeag tagaaagtte aateggaggg tgtggegaac caaacegatg egcaceacaa
                                                                    2400
atggaatggg atccaaatct agaatgttct tgtagcgacg aggagaaatg tgcgcatcat
                                                                    2460
                                                                    2520
tcccatcatt tctctctcga tattgatgtt ggatgtactg atttaaatga aaatctaggt
atatgggtta tatttaaaat taaaacgcag aacggttatg caaaattagg aaatttagag
                                                                    2580
tttctcgaag agaaaccatt aataggggaa gcgttagctc gtgttaagcg agtggagaaa
                                                                    2640
aaatggaaag acaaacgtga aaaattagaa tttgaaacga atatagtcta caacgaggca
                                                                    2700
aaagaagctg tggatgcact attcgtaaat tcacaatatg atagattgca agctgataca
                                                                    2760
aatatogoaa tgattoatgo ggoggataao aaagttoata aaattogoga ggogtacoto
                                                                    2820
ccagagttat ctgtgatacc aggtgtaaat gcgaccgttt ttgaagaatt agaagagcgt
                                                                    2880
atttttacag cattctccct ttacgatgca agaaatgtga taaaaaatgg ggatttcaat
                                                                    2940
aatggattat cttgttggaa tgtgaagggc caagtagatg tagaccaaaa tgaccatcgt
                                                                    3000
tetgteettg ttatteeagg atgggaateg gaagtateae aagaagttea tgtatgteea
                                                                    3060
gatcgtggat acattetteg tgttacggeg tacaaagaag gatatggaga aggetgegta
                                                                    3120
acaatccatg agattgataa tcatacagac gaactgaaat ttaaaaaactg ctttgaagag
                                                                    3180
gaagtatete taaataatge ggtgacatgt gatgagtata etacaaatea agaagtagga
                                                                    3240
gggtatgcgg atgtacgtca atccaataat cgtggatcta atgaggccta tgtaaatcct
                                                                    3300
acttccacat caactgatta tgcatcgctc tacgaggaag agtcgtatac gaatgaacag
                                                                    3360
acatataatt cttgtgaatc taacagaggg tatggtaatc aaatgccatt accgtctggc
                                                                    3420
tatgtgacaa aagaattaga atattttcca gagacagata aagtatggat tgagattgga
                                                                    3480
gaaacagaag gaacattcat cgtagacagt gtggaattac tccttatgga ggaataa
                                                                    3537
<210> SEO ID NO 8
<211> LENGTH: 1178
<212> TYPE: PRT
<213> ORGANISM: Bacillus thuringienses
<220> FEATURE:
<221> NAME/KEY: MISC FEATURE
```

<400> SEQUENCE: 8

Met Asp Asn Asn Ile Lys Asn Gln Cys Ile Pro Tyr Asn Cys Leu Asn 1 $$ 10 $$ 15

<222> LOCATION: (1)..(1178)

<223> OTHER INFORMATION: The amino acid sequence of the TIC2613 pesticidal protein.

-continued

_	_	~ 7		~-		_	~ 7	~ 7	~ 7	~ 7		m1	m1	-	_
Asp	Pro	Glu	20	GIu	11e	Leu	GIY	G1u 25	GIU	GIY	lle	Thr	Thr 30	ser	Asn
Glu	Asn	Leu 35	Glu	Phe	Phe	Leu	Ser 40	Leu	Thr	Lys	Phe	Val 45	Leu	Asn	Arg
Phe	Val 50	Pro	Gly	Gly	Ala	Tyr 55	Val	Ala	Gly	Leu	Phe 60	Asp	Val	Phe	Trp
Gly 65	Trp	Leu	Lys	Pro	Ser 70	Asp	Trp	Ser	Ala	Ile 75	Leu	Glu	Gln	Ile	Glu 80
Glu	Leu	Ile	Asn	Gln 85	Lys	Ile	Glu	Thr	Phe 90	Ala	Arg	Asn	Gln	Ala 95	Ile
Ser	Arg	Leu	Glu 100	Gly	Leu	Ser	Asn	Leu 105	Tyr	Glu	Ile	Tyr	Ala 110	Asp	Thr
Phe	Lys	Glu 115	Trp	Glu	Lys	Asp	Pro 120	Thr	Asn	Pro	Ala	Leu 125	Arg	Glu	Glu
Met	Arg 130	Thr	Gln	Phe	Asn	Asp 135	Met	Asn	Ser	Ser	Phe 140	Val	Thr	Ala	Met
Pro 145	Leu	Phe	Ser	Val	Gln 150	Asn	Phe	Glu	Val	Pro 155	Leu	Leu	Ala	Val	Tyr 160
Ala	Gln	Ala	Ala	Asn 165	Leu	His	Leu	Ser	Val 170	Leu	Arg	Asp	Val	Ser 175	Val
Phe	Gly	Gln	Asn 180	Trp	Gly	Phe	Asp	Ser 185	Ala	Thr	Val	Asn	Ser 190	Arg	Tyr
Asn	Asp	Leu 195	Val	Arg	Asn	Ile	Arg 200	Thr	Tyr	Thr	Asn	Tyr 205	Val	Val	Arg
Trp	Tyr 210	Asn	Thr	Gly	Leu	Ala 215	Arg	Leu	Arg	Gly	Thr 220	Thr	Tyr	Gln	Asp
Trp 225	Leu	Asn	Tyr	His	Arg 230	Phe	Arg	Arg	Glu	Leu 235	Thr	Ile	Thr	Ala	Leu 240
Asp	Ile	Ile	Thr	Ile 245	Phe	Pro	His	Tyr	Asp 250	Asn	Lys	Met	Tyr	Pro 255	Ile
Gln	Pro	His	Phe 260	Gln	Leu	Thr	Arg	Glu 265	Ile	Tyr	Thr	Asp	Pro 270	Leu	Ile
Asn	Phe	Asn 275	Pro	Ala	Leu	Gln	Ser 280	Val	Ala	Gln	Leu	Pro 285	Leu	Phe	Asn
Glu	Met 290	Glu	Asn	Ser	Thr	Ile 295	Arg	Ser	Pro	His	Leu 300	Val	Asp	Phe	Leu
Asn 305	Arg	Leu	Thr	Ile	Tyr 310	Thr	Aap	Trp	Tyr	Ser 315	Leu	Gly	Arg	His	Tyr 320
Tyr	Trp	Gly	Gly	His 325	Gln	Ile	Val	Ser	Arg 330	Gln	Thr	Gly	Ser	Thr 335	Ser
Thr	Ile	Thr	Phe 340	Pro	Ile	Tyr	Gly	Arg 345	Glu	Ala	Asn	Gln	Glu 350	Ala	Pro
Arg	Thr	Tyr 355	Asn	Phe	Ser	Gln	Pro 360	Val	Phe	Arg	Thr	Leu 365	Ser	Asn	Pro
Thr	Leu 370	Thr	Arg	Leu	Met	Gln 375	Pro	Trp	Pro	Ala	Pro 380	Ala	Phe	Gln	Leu
Arg 385	Arg	Leu	Glu	Gly	Val 390	Glu	Phe	Gln	Thr	Thr 395	Thr	Gly	Asn	Phe	Thr 400
Tyr	Arg	Gly	Arg	Gly 405	Thr	Val	Asp	Ser	Phe 410	Asp	Glu	Leu	Pro	Pro 415	Asp
Asp	Thr	Ser	Val 420	Pro	Ala	Arg	Glu	Gly 425	Tyr	Ser	His	Arg	Leu 430	Сла	His

Ala	Thr	Phe 435	Ile	Arg	Lys	Ser	Gly 440	Thr	Pro	Tyr	Leu	Thr 445	Thr	Gly	Val
Thr	Leu 450	Ser	Trp	Thr	His	Asn 455	Ser	Asn	Thr	Pro	Thr 460	Asn	Ile	Ile	Tyr
Pro 465	Asp	Lys	Ile	Thr	Gln 470	Val	Pro	Leu	Val	Lys 475	Ala	Ser	Asn	Leu	His 480
Ser	Ser	Ala	Phe	Val 485	Leu	Lys	Gly	Pro	Gly 490	Phe	Thr	Gly	Gly	Asp 495	Ile
Leu	Gly	Arg	Thr 500	Ser	Val	Gly	Asn	Ile 505	Ala	Asp	Ile	Gln	Met 510	Asn	Ile
Thr	Ala	Pro 515	Leu	Ser	Gln	Arg	Tyr 520	Arg	Val	Arg	Ile	Arg 525	Tyr	Ala	Ser
Thr	Thr 530	Asn	Leu	Gln	Phe	His 535	Thr	Thr	Ile	Asn	Gly 540	Arg	Ala	Val	Asn
Gln 545	Ala	Asn	Phe	Pro	Ala 550	Thr	Met	Asn	Arg	Val 555	Glu	Asp	Leu	Glu	Tyr 560
Asn	Ser	Phe	Arg	Thr 565	Ile	Ser	Phe	Gly	Thr 570	Pro	Phe	Asn	Phe	Leu 575	Asp
Ala	Gln	Ser	Thr 580	Phe	Arg	Leu	Gly	Val 585	Trp	Ser	Phe	Ser	Ser 590	Gly	Thr
Val	Leu	Ile 595	Asp	Arg	Ile	Glu	Val 600	Val	Pro	Met	Glu	Val 605	Thr	Phe	Glu
Ala	Glu 610	Ser	Asp	Leu	Glu	Arg 615	Ala	Gln	Lys	Ala	Val 620	Asn	Ala	Leu	Phe
Thr 625	Ser	Ile	Asn	Gln	630 Lys	Gly	Leu	Lys	Thr	Asp 635	Val	Thr	Asp	Tyr	His 640
Ile	Asp	Gln	Val	Ser 645	Asn	Leu	Val	Glu	Сув 650	Leu	Ser	Asp	Glu	Phe 655	CAa
Leu	Asp	Glu	660 660	Arg	Glu	Leu	Phe	Glu 665	ГÀа	Val	ГÀа	Tyr	Ala 670	ГÀа	Arg
Leu	Ser	Asp 675	Glu	Arg	Asn	Leu	Leu 680	Ala	Asp	Pro	Asn	Phe 685	Thr	Ser	Ile
Asn	Gly 690	Gln	Leu	Asp	Arg	Gly 695	Trp	Arg	Gly	Ser	Thr 700	Asp	Ile	Thr	Ile
Gln 705	Gly	Gly	Asp	Asp	Val 710	Phe	Lys	Glu	Asn	Tyr 715	Val	Thr	Leu	Ser	Gly 720
Thr	Leu	Asp	Glu	Сув 725	Tyr	Pro	Thr	Tyr	Leu 730	Tyr	Gln	ГÀа	Ile	Asp 735	Glu
Ser	ГЛа	Leu	Lys 740	Ala	Tyr	Thr	Arg	Tyr 745	Glu	Leu	Arg	Gly	Tyr 750	Ile	Glu
Asp	Ser	Gln 755	Asp	Leu	Glu	Val	Tyr 760	Leu	Ile	Arg	Tyr	Asn 765	Thr	ГÀа	His
Glu	Thr 770	Leu	Asn	Val	Pro	Gly 775	Thr	Gly	Gly	Leu	Trp 780	Pro	Leu	Ala	Val
Glu 785	Ser	Ser	Ile	Gly	Gly 790	Сув	Gly	Glu	Pro	Asn 795	Arg	Cys	Ala	Pro	Gln 800
Met	Glu	Trp	Asp	Pro 805	Asn	Leu	Glu	СЛа	Ser 810	Сув	Ser	Asp	Glu	Glu 815	ГÀз
CÀa	Ala	His	His 820	Ser	His	His	Phe	Ser 825	Leu	Asp	Ile	Asp	Val 830	Gly	CAa
Thr	Asp	Leu 835	Asn	Glu	Asn	Leu	Gly 840	Ile	Trp	Val	Ile	Phe 845	Lys	Ile	ГХа
Thr	Gln	Asn	Gly	Tyr	Ala	Lys	Leu	Gly	Asn	Leu	Glu	Phe	Leu	Glu	Glu

-continued

										-	-cor	ıtır	ıued	
85	50				855					860)			
Lys Pr 865	ro Leu	Ile	Gly	Glu 870	Ala	Leu	Ala	Arg	Val 875	Lys	Arg	Val	. Glu	880 Lys
Lys Tr	cb ràa	Asp	Lys 885	Arg	Glu	Lys	Leu	Glu 890	Phe	Glu	Thr	Asn	11e 895	
Tyr As	sn Glu	Ala 900	ГÀа	Glu	Ala		Asp 905	Ala	Leu	Phe	· Val	Asr 910		Gln
Tyr As	sp Arg 915	Leu	Gln	Ala	-	Thr . 920	Asn	Ile	Ala	Met	11e 925		. Ala	Ala
	en Lys	Val	His	Lys	Ile 935	Arg	Glu	Ala	Tyr	Leu 940		Glu	ı Leu	Ser
Val II 945	le Pro	Gly	Val	Asn 950	Ala	Thr	Val	Phe	Glu 955	Glu	Leu	. Glu	ı Glu	Arg 960
Ile Ph	ne Thr	Ala	Phe 965	Ser	Leu	Tyr	Asp	Ala 970	Arg	Asn	ı Val	Ile	975	
Gly As	sp Phe	Asn 980	Asn	Gly	Leu		Cys 985	Trp	Asn	Val	. Lys	Gly 990		Val
Asp Va	al Asp 995	Gln	Asn	Asp		Arg 1000		. Val	L Le	ı Va		e F 05	ro G	ly Trp
	er Glu 010	ı Val	. Ser	Glr	Glu 101		1 H:	is Va	al Cy		ro .020	Asp	Arg	Gly
	le Let)25	ı Arç	y Val	Thr	Ala 103		r Ly	/s G	Lu G		'yr .035	Gly	Glu	Gly
	al Thi	r Ile	e His	Glu	lle 104		p As	en H	is Tl		.050	Glu	Leu	Lys
	/s Asi)55	n Cys	Phe	Glu	Glu 106		u Vá	al Se	er Le		sn .065	Asn	Ala	Val
	/s As <u>r</u>)70	Glu	ı Tyr	Thr	Thr 107		n G	ln G	Lu Va		ly .080	Gly	Tyr	Ala
_	al Arç 085	g Glr	n Ser	Asr	109		g GI	Ly Se	er As		lu .095	Ala	Tyr	Val
	o Thi	r Sei	Thr	Ser	Thr 110		рΤχ	/r A	La Se		eu .110	Tyr	Glu	Glu
	er Tyn 115	r Thi	Asr.	Glu	112		r Ty	/r As	sn Se		'ys .125	Glu	Ser	Asn
	Ly Ty: L30	r Gly	/ Asr	Glr	113	Pr 5	o Le	eu Pi	co Se	-	1y .140	Tyr	Val	Thr
	lu Lei 145	ı Glu	ı Tyr	Ph∈	Pro 115		u Tł	ır As	ab ri		al 155	Trp	Ile	Glu
	ly Glu 160	ı Thı	Glu	Gly	7 Thr 116		e II	Le Va	al As	_	er 170	Val	Glu	Leu
	eu Met L75	Glu	ı Glu	L										
<211><212><213><220>		H: 35 DNA ISM: RE: INFO L3PL where	Arti Arti RMAT pest	'ION: icid n ad	A s lal p lditi	rote onal	in o	desiq anine	gned e coo	for don	exp is i	ress	ion	ng a in a plant immediately

<400> SEQUENCE: 9

				-contir	nued	
gtcgagatcc	tcggcgagga	gggcataacg	acgagcaacg	agaaccttga	gttcttcctc	120
agcctcacga	agttcgtcct	gaaccgcttc	gtgccgggcg	gagcctacgt	ggctggcctg	180
ttcgatgtgt	tctggggatg	gctcaagcca	agcgactggt	ccgcgattct	ggagcagatc	240
gaggaactca	tcaaccagaa	gatcgagaca	ttcgcccgca	accaggccat	cagccgcctg	300
gagggcctct	cgaacctcta	cgaaatctac	gccgatacgt	tcaaggagtg	ggagaaggat	360
ccgacgaacc	cggccttgcg	cgaggagatg	aggacgcaat	tcaacgacat	gaactccagc	420
ttegtcaceg	ccatgccgct	gttctccgtc	cagaacttcg	aggtgccctt	gctcgccgtg	480
tacgcgcaag	ctgcgaactt	acatcttagc	gteeteegeg	acgtcagcgt	cttcggccag	540
aactggggat	tcgattccgc	gacggtgaac	tcacggtaca	atgatctcgt	gcggaacatc	600
cggacctaca	ccaattacgt	cgtgcgctgg	tacaacacgg	gattggcgcg	tctgcgcggc	660
actacctacc	aggactggct	caactaccac	cggttccgcc	gcgaactcac	aatcacagcg	720
ctggacatca	ttaccatctt	cccgcactac	gacaacaaga	tgtacccaat	ccagcctcac	780
ttccagctta	cccgtgagat	ctacacggac	ccgctcatca	acttcaatcc	cgcactgcaa	840
tcagtagccc	aattgccact	cttcaacgag	atggagaact	cgacaatccg	aagccctcac	900
ctcgtggact	tcctcaaccg	cctgaccatc	tacacggatt	ggtactctct	tggtcggcac	960
tactattggg	gegggeacea	aatcgtgtcc	aggcagaccg	gctctacctc	taccataacc	1020
ttcccgatct	atggccggga	ggccaaccag	gaggeteega	ggacttacaa	cttcagtcag	1080
ccagtgttcc	gcacactctc	caacccgact	ctcactcgtt	tgatgcagcc	ctggcccgct	1140
cccgcgtttc	agctcagaag	attggagggc	gtggagttcc	aaacaacgac	gggcaacttc	1200
acctaccgtg	gccgtgggac	ggtggacagt	ttcgacgagt	tgcctccgga	cgacaccagc	1260
gtgcctgcaa	gggaaggcta	ctcgcacagg	ctgtgccacg	cgacgttcat	ccgcaagtct	1320
gggacaccct	acctgacaac	cggcgtcact	ctctcctgga	cccacaacag	caacacaccc	1380
accaacataa	tctaccctga	caagataaca	caagtgccgc	tggtgaaggc	ttcgaacctc	1440
cattcctccg	ccttcgtcct	caagggtccg	ggcttcaccg	gcggcgacat	cctgggtcgc	1500
acgtcggtcg	gcaacatcgc	ggacattcag	atgaacatta	ccgcacctct	gtcccagcgc	1560
tacagagtgc	gtatccgcta	cgcgagtacg	accaacctcc	aattccacac	tacgatcaat	1620
gggagggcgg	ttaatcaggc	caacttcccg	gccacgatga	accgggtcga	agacctggag	1680
tacaactcgt	ttcggaccat	ctctttcggc	acgccgttca	acttcctaga	cgcccagtca	1740
acctttcggc	tgggagtttg	gagcttcagc	agcggcacag	tcctcatcga	ccgaatagag	1800
gtggttccga	tggaggtcac	gttcgaggcg	gagtcggacc	tggagcgagc	gcagaaggct	1860
gtaaatgcgt	tgttcacgag	cattaaccag	aagggcctca	agaccgatgt	cacagactac	1920
cacatcgacc	aagtgtcgaa	cctggtggag	tgtctgtcgg	atgagttctg	tcttgacgag	1980

gtcgagatcc	teggegagga	gggcataacg	acgagcaacg	agaaccttga	gttetteete	120
agcctcacga	agttcgtcct	gaaccgcttc	gtgccgggcg	gagcctacgt	ggctggcctg	180
ttcgatgtgt	tctggggatg	gctcaagcca	agcgactggt	ccgcgattct	ggagcagatc	240
gaggaactca	tcaaccagaa	gatcgagaca	ttcgcccgca	accaggccat	cagccgcctg	300
gagggcctct	cgaacctcta	cgaaatctac	gccgatacgt	tcaaggagtg	ggagaaggat	360
ccgacgaacc	cggccttgcg	cgaggagatg	aggacgcaat	tcaacgacat	gaactccagc	420
ttcgtcaccg	ccatgccgct	gttctccgtc	cagaacttcg	aggtgccctt	gctcgccgtg	480
tacgcgcaag	ctgcgaactt	acatcttagc	gtcctccgcg	acgtcagcgt	cttcggccag	540
aactggggat	tegatteege	gacggtgaac	tcacggtaca	atgatctcgt	gcggaacatc	600
cggacctaca	ccaattacgt	cgtgcgctgg	tacaacacgg	gattggcgcg	tctgcgcggc	660
actacctacc	aggactggct	caactaccac	eggtteegee	gcgaactcac	aatcacagcg	720
ctggacatca	ttaccatctt	cccgcactac	gacaacaaga	tgtacccaat	ccagcctcac	780
ttccagctta	cccgtgagat	ctacacggac	ccgctcatca	acttcaatcc	cgcactgcaa	840
tcagtagccc	aattgccact	cttcaacgag	atggagaact	cgacaatccg	aagccctcac	900
ctcgtggact	tcctcaaccg	cctgaccatc	tacacggatt	ggtactctct	tggtcggcac	960
tactattggg	gcgggcacca	aatcgtgtcc	aggcagaccg	gctctacctc	taccataacc	1020
ttcccgatct	atggccggga	ggccaaccag	gaggctccga	ggacttacaa	cttcagtcag	1080
ccagtgttcc	gcacactctc	caacccgact	ctcactcgtt	tgatgcagcc	ctggcccgct	1140
cccgcgtttc	agctcagaag	attggagggc	gtggagttcc	aaacaacgac	gggcaacttc	1200
acctaccgtg	gccgtgggac	ggtggacagt	ttcgacgagt	tgcctccgga	cgacaccagc	1260
gtgcctgcaa	gggaaggcta	ctcgcacagg	ctgtgccacg	cgacgttcat	ccgcaagtct	1320
gggacaccct	acctgacaac	cggcgtcact	ctctcctgga	cccacaacag	caacacaccc	1380
accaacataa	tctaccctga	caagataaca	caagtgeege	tggtgaaggc	ttcgaacctc	1440
catteeteeg	ccttcgtcct	caagggtccg	ggcttcaccg	gcggcgacat	cctgggtcgc	1500
acgtcggtcg	gcaacatcgc	ggacattcag	atgaacatta	ccgcacctct	gtcccagcgc	1560
tacagagtgc	gtatccgcta	cgcgagtacg	accaacctcc	aattccacac	tacgatcaat	1620
gggagggcgg	ttaatcaggc	caacttcccg	gccacgatga	accgggtcga	agacctggag	1680
tacaactcgt	ttcggaccat	ctctttcggc	acgccgttca	acttcctaga	cgcccagtca	1740
acctttcggc	tgggagtttg	gagetteage	agcggcacag	tcctcatcga	ccgaatagag	1800
gtggttccga	tggaggtcac	gttcgaggcg	gagtcggacc	tggagcgagc	gcagaaggct	1860
gtaaatgcgt	tgttcacgag	cattaaccag	aagggcctca	agaccgatgt	cacagactac	1920
cacatcgacc	aagtgtcgaa	cctggtggag	tgtctgtcgg	atgagttctg	tcttgacgag	1980
aagcgggagc	tgttcgagaa	ggtgaagtat	gctaagcggc	tgagcgacga	gcggaacttg	2040
ttggctgacc	cgaacttcac	cagcatcaac	ggacageteg	accgtgggtg	gcgaggttcc	2100
accgacatca	cgatacaggg	cggagacgat	gtgttcaagg	agaactatgt	gaccctctca	2160
ggaacactgg	atgagtgcta	cccgacctat	ctctaccaga	agatcgacga	gagcaagctc	2220
aaggcttaca	cgcgctacga	actccgtggc	tacatcgaag	actcccagga	tcttgaggtg	2280
tacctcatac	gctacaacac	aaagcacgag	acgctcaacg	ttcctggcac	cggtggtctt	2340
tggcccttgg	ccgtggagag	tagcatcggc	gggtgcggtg	agccaaaccg	atgcgcgcca	2400

-continued

cagatggaat	gggatccgaa	cctagagtgc	tcctgctcag	acgaggagaa	gtgcgcccac	2460
cactcccacc	acttctcgct	cgacattgac	gttggctgca	cggatctcaa	cgagaaccta	2520
ggaatctggg	tgatcttcaa	gattaagacc	cagaacggct	acgccaagct	cgggaatctg	2580
gagtttcttg	aggagaagcc	gctgatcggc	gaggccctcg	cgcgcgtgaa	gcgagtcgag	2640
aagaagtgga	aagacaagcg	ggagaagcta	gagtttgaaa	cgaacattgt	ttacaacgag	2700
gcaaaggaag	ccgtggacgc	tctgttcgta	aacagtcagt	acgaccgtct	ccaggccgac	2760
acgaacatcg	caatgataca	cgcggcggat	aacaaggtgc	acaagattcg	ggaggcttac	2820
ctgcccgagc	tgtcggtcat	cccaggcgta	aacgctaccg	tgttcgagga	gctggaggaa	2880
cggatcttca	ccgcgttctc	cctctatgac	gcaaggaacg	tcatcaagaa	cggcgacttc	2940
aacaacggcc	tgagctgctg	gaacgtgaag	ggccaagtgg	acgtcgatca	gaacgatcac	3000
cgctccgttc	tggtcattcc	agggtgggag	tccgaggtga	gccaagaggt	ccatgtgtgc	3060
ccggaccgtg	gctacatcct	tegggtgace	gcgtacaagg	agggctacgg	cgaaggctgc	3120
gtgaccatac	acgagatcga	caaccacacc	gacgagctta	agttcaagaa	ctgcttcgag	3180
gaggaggtgt	cactgaacaa	cgccgtgacc	tgcgacgagt	acacgaccaa	tcaggaggtc	3240
ggcggctacg	ccgacgtccg	ccagtcgaac	aatcgaggca	gcaacgaggc	gtacgtgaac	3300
ccaacctcca	cctcgacgga	ctacgccagc	ctctacgagg	aggagtccta	cacaaacgag	3360
cagacctaca	actcgtgcga	gagcaaccga	ggttacggga	accagatgcc	gctaccgtcc	3420
gggtacgtga	cgaaggagct	ggagtatttc	ccagagaccg	acaaggtgtg	gatcgagatc	3480
ggcgagacag	agggcacgtt	catcgtggac	agcgtcgagc	tgctgttgat	ggaggagtga	3540

<210> SEQ ID NO 10

<400> SEQUENCE: 10

Met Ala Asp Asn Asn Ile Lys Asn Gln Cys Ile Pro Tyr Asn Cys Leu 1 $$ 5 $$ 10 $$ 15

Asn Glu Asn Leu Glu Phe Phe Leu Ser Leu Thr Lys Phe Val Leu Asn 35 40 45

Arg Phe Val Pro Gly Gly Ala Tyr Val Ala Gly Leu Phe Asp Val Phe 50 55 60

Trp Gly Trp Leu Lys Pro Ser Asp Trp Ser Ala Ile Leu Glu Gln Ile 65 $$ 70 $$ 75 $$ 80

Glu Glu Leu Ile Asn Gln Lys Ile Glu Thr Phe Ala Arg Asn Gln Ala \$85\$ 90 95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Glu Ile Tyr Ala Asp \$100\$

Thr Phe Lys Glu Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu 115 120 125

Glu Met Arg Thr Gln Phe Asn Asp Met Asn Ser Ser Phe Val Thr Ala 130 $$135\$

Met Pro Leu Phe Ser Val Gln Asn Phe Glu Val Pro Leu Leu Ala Val

<211> LENGTH: 1179

<212> TYPE: PRT

<213> ORGANISM: Artificial

<220> FEATURE:

<223> OTHER INFORMATION: The amino acid sequence of TIC2613PL encoded by a synthetic coding sequence designed for expression in a plant cell (SEQ ID NO:9), and wherein an additional alanine amino acid is inserted immediately following the initiating methionine.

-continued

_															
145					150					155					160
Tyr	Ala	Gln	Ala	Ala 165	Asn	Leu	His	Leu	Ser 170	Val	Leu	Arg	Asp	Val 175	Ser
Val	Phe	Gly	Gln 180	Asn	Trp	Gly	Phe	Asp 185	Ser	Ala	Thr	Val	Asn 190	Ser	Arg
Tyr	Asn	Asp 195	Leu	Val	Arg	Asn	Ile 200	Arg	Thr	Tyr	Thr	Asn 205	Tyr	Val	Val
Arg	Trp 210	Tyr	Asn	Thr	Gly	Leu 215	Ala	Arg	Leu	Arg	Gly 220	Thr	Thr	Tyr	Gln
Asp 225	Trp	Leu	Asn	Tyr	His 230	Arg	Phe	Arg	Arg	Glu 235	Leu	Thr	Ile	Thr	Ala 240
Leu	Asp	Ile	Ile	Thr 245	Ile	Phe	Pro	His	Tyr 250	Asp	Asn	Lys	Met	Tyr 255	Pro
Ile	Gln	Pro	His 260	Phe	Gln	Leu	Thr	Arg 265	Glu	Ile	Tyr	Thr	Asp 270	Pro	Leu
Ile	Asn	Phe 275	Asn	Pro	Ala	Leu	Gln 280	Ser	Val	Ala	Gln	Leu 285	Pro	Leu	Phe
Asn	Glu 290	Met	Glu	Asn	Ser	Thr 295	Ile	Arg	Ser	Pro	His 300	Leu	Val	Asp	Phe
Leu 305	Asn	Arg	Leu	Thr	Ile 310	Tyr	Thr	Asp	Trp	Tyr 315	Ser	Leu	Gly	Arg	His 320
Tyr	Tyr	Trp	Gly	Gly 325	His	Gln	Ile	Val	Ser 330	Arg	Gln	Thr	Gly	Ser 335	Thr
Ser	Thr	Ile	Thr 340	Phe	Pro	Ile	Tyr	Gly 345	Arg	Glu	Ala	Asn	Gln 350	Glu	Ala
Pro	Arg	Thr 355	Tyr	Asn	Phe	Ser	Gln 360	Pro	Val	Phe	Arg	Thr 365	Leu	Ser	Asn
Pro	Thr 370	Leu	Thr	Arg	Leu	Met 375	Gln	Pro	Trp	Pro	Ala 380	Pro	Ala	Phe	Gln
Leu 385	Arg	Arg	Leu	Glu	Gly 390	Val	Glu	Phe	Gln	Thr 395	Thr	Thr	Gly	Asn	Phe 400
Thr	Tyr	Arg	Gly	Arg 405	Gly	Thr	Val	Asp	Ser 410	Phe	Asp	Glu	Leu	Pro 415	Pro
Asp	Asp	Thr	Ser 420	Val	Pro	Ala	Arg	Glu 425	Gly	Tyr	Ser	His	Arg 430	Leu	СЛа
His	Ala	Thr 435	Phe	Ile	Arg	ГÀЗ	Ser 440	Gly	Thr	Pro	Tyr	Leu 445	Thr	Thr	Gly
Val	Thr 450	Leu	Ser	Trp	Thr	His 455	Asn	Ser	Asn	Thr	Pro 460	Thr	Asn	Ile	Ile
Tyr 465	Pro	Asp	Lys	Ile	Thr 470	Gln	Val	Pro	Leu	Val 475	Lys	Ala	Ser	Asn	Leu 480
His	Ser	Ser	Ala	Phe 485	Val	Leu	Lys	Gly	Pro 490	Gly	Phe	Thr	Gly	Gly 495	Asp
Ile	Leu	Gly	Arg 500	Thr	Ser	Val	Gly	Asn 505	Ile	Ala	Asp	Ile	Gln 510	Met	Asn
Ile	Thr	Ala 515	Pro	Leu	Ser	Gln	Arg 520	Tyr	Arg	Val	Arg	Ile 525	Arg	Tyr	Ala
Ser	Thr 530	Thr	Asn	Leu	Gln	Phe 535	His	Thr	Thr	Ile	Asn 540	Gly	Arg	Ala	Val
Asn 545	Gln	Ala	Asn	Phe	Pro 550	Ala	Thr	Met	Asn	Arg 555	Val	Glu	Asp	Leu	Glu 560
Tyr	Asn	Ser	Phe	Arg 565	Thr	Ile	Ser	Phe	Gly 570	Thr	Pro	Phe	Asn	Phe 575	Leu

Asp	Ala	Gln	Ser 580	Thr	Phe	Arg	Leu	Gly 585	Val	Trp	Ser	Phe	Ser 590	Ser	Gly
Thr	Val	Leu 595	Ile	Asp	Arg	Ile	Glu 600	Val	Val	Pro	Met	Glu 605	Val	Thr	Phe
Glu	Ala 610	Glu	Ser	Asp	Leu	Glu 615	Arg	Ala	Gln	Lys	Ala 620	Val	Asn	Ala	Leu
Phe 625	Thr	Ser	Ile	Asn	Gln 630	Lys	Gly	Leu	Lys	Thr 635	Asp	Val	Thr	Asp	Tyr 640
His	Ile	Asp	Gln	Val 645	Ser	Asn	Leu	Val	Glu 650	Сув	Leu	Ser	Asp	Glu 655	Phe
Cys	Leu	Asp	Glu 660	Lys	Arg	Glu	Leu	Phe 665	Glu	Lys	Val	ГÀа	Tyr 670	Ala	ГХа
Arg	Leu	Ser 675	Asp	Glu	Arg	Asn	Leu 680	Leu	Ala	Asp	Pro	Asn 685	Phe	Thr	Ser
Ile	Asn 690	Gly	Gln	Leu	Asp	Arg 695	Gly	Trp	Arg	Gly	Ser 700	Thr	Asp	Ile	Thr
Ile 705	Gln	Gly	Gly	Asp	Asp 710	Val	Phe	Lys	Glu	Asn 715	Tyr	Val	Thr	Leu	Ser 720
Gly	Thr	Leu	Asp	Glu 725	CAa	Tyr	Pro	Thr	Tyr 730	Leu	Tyr	Gln	Lys	Ile 735	Asp
Glu	Ser	Lys	Leu 740	Lys	Ala	Tyr	Thr	Arg 745	Tyr	Glu	Leu	Arg	Gly 750	Tyr	Ile
Glu	Asp	Ser 755	Gln	Asp	Leu	Glu	Val 760	Tyr	Leu	Ile	Arg	Tyr 765	Asn	Thr	ГЛа
His	Glu 770	Thr	Leu	Asn	Val	Pro 775	Gly	Thr	Gly	Gly	Leu 780	Trp	Pro	Leu	Ala
Val 785	Glu	Ser	Ser	Ile	Gly 790	Gly	Cys	Gly	Glu	Pro 795	Asn	Arg	Cys	Ala	Pro 800
Gln	Met	Glu	Trp	Asp 805	Pro	Asn	Leu	Glu	Cys 810	Ser	CAa	Ser	Asp	Glu 815	Glu
Lys	Сув	Ala	His 820	His	Ser	His	His	Phe 825	Ser	Leu	Asp	Ile	Asp 830	Val	Gly
Cys	Thr	Asp 835	Leu	Asn	Glu	Asn	Leu 840	Gly	Ile	Trp	Val	Ile 845	Phe	Lys	Ile
Lys	Thr 850	Gln	Asn	Gly	Tyr	Ala 855	Lys	Leu	Gly	Asn	Leu 860	Glu	Phe	Leu	Glu
Glu 865	Lys	Pro	Leu	Ile	Gly 870	Glu	Ala	Leu	Ala	Arg 875	Val	Lys	Arg	Val	Glu 880
Lys	Lys	Trp	Lys	Asp 885	Lys	Arg	Glu	Lys	Leu 890	Glu	Phe	Glu	Thr	Asn 895	Ile
Val	Tyr	Asn	Glu 900	Ala	Lys	Glu	Ala	Val 905	Asp	Ala	Leu	Phe	Val 910	Asn	Ser
Gln	Tyr	Asp 915	Arg	Leu	Gln	Ala	Asp 920	Thr	Asn	Ile	Ala	Met 925	Ile	His	Ala
Ala	Asp 930	Asn	Lys	Val	His	Lys 935	Ile	Arg	Glu	Ala	Tyr 940	Leu	Pro	Glu	Leu
Ser 945	Val	Ile	Pro	Gly	Val 950	Asn	Ala	Thr	Val	Phe 955	Glu	Glu	Leu	Glu	Glu 960
Arg	Ile	Phe	Thr	Ala 965	Phe	Ser	Leu	Tyr	Asp 970	Ala	Arg	Asn	Val	Ile 975	Lys
Asn	Gly	Asp	Phe 980	Asn	Asn	Gly	Leu	Ser 985	Сла	Trp	Asn	Val	Lys	Gly	Gln

-continued

Val		Val 1 995	Asp (Gln A	Asn A		is 1	Arg :	Ser V	Jal 1		al :	Ile I	Pro Gly	7				
Trp	Glu 1010		Glu	Val	Ser	Gln 1015	Glu	Val	His	Val	Cys 1020	Pro	Asp	Arg					
Gly	Tyr 1025		Leu	Arg	Val	Thr 1030	Ala	Tyr	Lys	Glu	Gly 1035	Tyr	Gly	Glu					
Gly	Cys 1040		Thr	Ile	His	Glu 1045	Ile	Asp	Asn	His	Thr 1050	Asp	Glu	Leu					
Lys	Phe 1055		Asn	Cys	Phe	Glu 1060	Glu	Glu	Val	Ser	Leu 1065	Asn	Asn	Ala					
Val	Thr 1070	•	Asp	Glu	Tyr	Thr 1075	Thr	Asn	Gln	Glu	Val 1080	Gly	Gly	Tyr					
Ala	Asp 1085		Arg	Gln	Ser	Asn 1090	Asn	Arg	Gly	Ser	Asn 1095	Glu	Ala	Tyr					
Val	Asn 1100		Thr	Ser	Thr	Ser 1105	Thr	Asp	Tyr	Ala	Ser 1110	Leu	Tyr	Glu					
Glu	Glu 1115		Tyr	Thr	Asn	Glu 1120	Gln	Thr	Tyr	Asn	Ser 1125	Cys	Glu	Ser					
Asn	Arg 1130		Tyr	Gly	Asn	Gln 1135	Met	Pro	Leu	Pro	Ser 1140	Gly	Tyr	Val					
Thr	Lys 1145		Leu	Glu	Tyr	Phe 1150	Pro	Glu	Thr	Asp	Lys 1155	Val	Trp	Ile					
Glu	Ile 1160	•	Glu	Thr	Glu	Gly 1165	Thr	Phe	Ile	Val	Asp 1170	Ser	Val	Glu					
Leu	Leu 1175		Met	Glu	Glu														

What is claimed is:

- 1. A recombinant nucleic acid molecule comprising a heterologous promoter operably linked to a polynucleotide segment encoding a pesticidal protein or pesticidal fragment thereof, wherein:
 - a. said pesticidal protein comprises the amino acid sequence of SEQ ID NO: 10; or said pesticidal protein comprises an amino add sequence having at least 90%, or 95%, or about 100% amino add sequence identity to SEO ID NO:10; or
 - b. said polynucleotide segment hybridizes under stringent hybridization conditions to a polynucleotide having the nucleotide sequence of SEQ ID NO:9.
- 2. The recombinant nucleic acid molecule of claim 1, wherein:
 - a. the recombinant nucleic acid molecule comprises a 50 sequence that functions to express the pesticidal protein in a plant; or
 - b. the recombinant nucleic acid molecule is expressed in a plant cell to produce a pesticidally effective amount of the pesticidal protein; or
 - c. the recombinant nucleic acid molecule is in operable linkage with a vector, and said vector is selected from the group consisting of a plasmid, phagemid, bacmid, cosmid, and a bacterial or yeast artificial chromosome.
- **3**. A host cell comprising the recombinant nucleic acid 60 molecule of claim **1**, wherein said host cell is selected from the group consisting of a bacterial and a plant cell.
- 4. The host cell of claim 3, wherein the bacterial host cell is from a genus of bacteria selected from the group consisting of: *Agrobacterium, Rhizobium, Bacillus, Brevibacillus,* 65 *Escherichia, Pseudomonas, Klebsiella, Pantoea*, and *Erwinia*.

- 5. The host cell of claim 4, wherein the *Bacillus* species is *Bacillus cereus* or *Bacillus thuringiensis*, said *Brevibacillus* is *Brevibacillus laterosperous*, and said *Escherichia* is *Escherichia coli*.
- **6**. The host cell of claim **3**, wherein said plant cell is a dicotyledonous or a monocotyledonous plant cell.
- 7. The host cell of claim 6, wherein said plant host cell is selected from the group consisting of an alfalfa, banana, barley, bean, broccoli, cabbage, brassica, carrot, cassava, castor, cauliflower, celery, chickpea, Chinese cabbage, citrus, coconut, coffee, corn, clover, cotton, a cucurbit, cucumber, Douglas fir, eggplant, eucalyptus, flax, garlic, grape, hops, leek, lettuce, Loblolly pine, millets, melons, nut, oat, olive, onion, ornamental, palm, pasture grass, pea, peanut, pepper, pigeonpea, pine, potato, poplar, pumpkin, Radiata pine, radish, rapeseed, rice, rootstocks, rye, safflower, shrub, sorghum, Southern pine, soybean, spinach, squash, strawberry, sugar beet, sugarcane, sunflower, sweet corn, sweet gum, sweet potato, switchgrass, tea, tobacco, tomato, triticale, turf grass, watermelon, and wheat plant cell.
- 8. The recombinant nucleic acid molecule of claim 1, wherein said protein exhibits activity against a Lepidopteran insect.
- 9. The recombinant nucleic acid molecule of claim 8, wherein said insect is selected from the group consisting of: Beet armyworm (Spodoptera exigua), Corn earworm (Helicoverpa zea), Cotton leaf worm (Alabama argillacea), European corn borer (Ostrinia nubilalis), Fall armyworm (Spodoptera frugiperda), Old World bollworm (Helicoverpa armigera), Oriental leaf worm (Spodoptera litura), Pink bollworm (Pectinophora gossypiella), Cry1Ac resistant Pink bollworm (Pectinophora gossypiella), Soybean looper

80

(Chrysodeixis includens), Southern armyworm (Spodoptera eridania), Southwestern corn borer (Diatraea grandiosella), Spotted bollworm (Earias vittella), Sugarcane borer (Diatraea saccharalis), Tobacco budworm (Heliothis virescens), and Velvet bean caterpillar (Anticarsia gemmatalis).

- 10. A plant or part thereof comprising the recombinant nucleic acid molecule of claim 1.
- 11. The plant or part thereof of claim 10, wherein said plant is a monocot plant or a dicot plant.
- 12. The plant of claim 10, wherein the plant is selected from the group consisting of an alfalfa, banana, barley, bean, broccoli, cabbage, brassica, carrot, cassava, castor, cauliflower, celery, chickpea, Chinese cabbage, citrus, coconut, coffee, corn, clover, cotton, a cucurbit, cucumber, Douglas fir, eggplant, eucalyptus, flax, garlic, grape, hops, leek, lettuce, Loblolly pine, millets, melons, nut, oat, olive, onion, ornamental, palm, pasture grass, pea, peanut, pepper, pigeon pea, pine, potato, poplar, pumpkin, Radiata pine, radish, rapeseed, rice, rootstocks, rye, safflower, shrub, sorghum, 20 Southern pine, soybean, spinach, squash, strawberry, sugar beet, sugarcane, sunflower, sweet corn, sweet gum, sweet potato, switchgrass, tea, tobacco, tomato, triticale, turf grass, watermelon, and wheat.
- 13. A seed of the plant of claim 10, wherein said seed 25 comprises said recombinant nucleic acid molecule.
- 14. An insect inhibitory composition comprising the recombinant nucleic acid molecule of claim 1.
- 15. The insect inhibitory composition of claim 14, further comprising a nucleotide sequence encoding at least one 30 other pesticidal agent that is different from said pesticidal protein.
- 16. The insect inhibitory composition of claim 15, wherein said at least one other pesticidal agent is selected from the group consisting of an insect inhibitory protein, an 35 insect inhibitory dsRNA molecule, and an ancillary protein.
- 17. The insect inhibitory composition of claim 16, wherein said at least one other pesticidal agent exhibits activity against one or more pest species of the orders Lepidoptera, Coleoptera, or Hemiptera.
- **18**. The insect inhibitory composition of claim **17**, wherein said at least one other pesticidal protein is selected from the group consisting of a Cry1A, Cry1Ab, Cry1Ac, Cry1A.105, Cry1Ae, Cry1B, Cry1C, Cry1C variants, Cry1D, Cry1E, Cry1F, Cry1A/F chimeras, Cry1G, Cry1H, 45 Cry1I, Cry1J, Cry1K, Cry1L, Cry2A, Cry2Ab, Cry2Ae, Cry3, Cry3A variants, Cry3B, Cry4B, Cry6, Cry7, Cry8, Cry9, Cry15, Cry34, Cry35, Cry43A, Cry43B, Cry51Aa1, ET29, ET33, ET34, ET35, ET66, ET70, TIC400, TIC407, TIC417, TIC431, TIC800, TIC807, TIC834, TIC853, 50 TIC900, TIC901, TIC1201, TIC1415, TIC2160, TIC3131, TIC836, TIC860, TIC867, TIC869, TIC1100, VIP3A, VIP3B, VIP3Ab, AXMI-AXMI-, AXMI-88, AXMI-97, AXMI-102, AXMI-112, AXMI-117, AXMI-100, AXMI-115, AXMI-113, and AXMI-005, AXMI134, AXMI-150, 55 AXMI-171, AXMI-184, AXMI-196, AXMI-204, AXMI-207, AXMI-209, AXMI-205, AXMI-218, AXMI-220, AXMI-221z, AXMI-222z, AXMI-223z, AXMI-224z and AXMI-225z, AXMI-238, AXMI-270, AXMI-279, AXMI-345, AXMI-335, AXMI-R1 and variants thereof, IP3 and 60 variants thereof, DIG-3, DIG-5, DIG-10, DIG-657 and a DIG-11 protein.
- 19. The insect inhibitory composition of claim 14, defined as comprising a plant cell that expresses said recombinant nucleic acid molecule.
- 20. A commodity product produced from the plant or part thereof of claim 10, wherein the commodity product com-

82

prises a detectable amount of said recombinant nucleic acid molecule or a pesticidal protein.

- 21. The commodity product of claim 20, selected from the group consisting of commodity corn bagged by a grain handler, corn flakes, corn cakes, corn flour, corn meal, corn syrup, corn oil, corn silage, corn starch, corn cereal, and the like, and corresponding sovbean, rice, wheat, sorghum, pigeon pea, peanut, fruit, melon, and vegetable commodity products including, where applicable, juices, concentrates, jams, jellies, marmalades, and other edible forms of such commodity products containing a detectable amount of such polynucleotides and or polypeptides of this application, whole or processed cotton seed, cotton oil, lint, seeds and plant parts processed for feed or food, fiber, paper, biomasses, and fuel products such as fuel derived from cotton oil or pellets derived from cotton gin waste, whole or processed soybean seed, soybean oil, soybean protein, soybean meal, soybean flour, soybean flakes, soybean bran, soybean milk, soybean cheese, soybean wine, animal feed comprising soybean, paper comprising soybean, cream comprising soybean, soybean biomass, and fuel products produced using soybean plants and soybean plant parts.
 - 22. A method of producing seed comprising:
 - a. planting a first seed according to claim 13;
 - b. growing a plant from the seed; and
 - c. harvesting seed from the plants, wherein said harvested seed comprises said recombinant nucleic acid molecule
- 23. A plant resistant to insect infestation, wherein the cells of said plant comprise the recombinant nucleic acid molecule of claim 1.
- **24**. A method for controlling a Lepidopteran species pest or pest infestation, said method comprising:
 - a. contacting the pest with an insecticidally effective amount of a pesticidal protein as set forth in SEQ ID NO: 10; or contacting the pest with an insecticidally effective amount of one or more pesticidal proteins comprising an amino acid sequence having at least 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:10.
- **25**. A method of detecting the presence of the recombinant nucleic acid molecule of claim 1 in a sample comprising plant genomic DNA, comprising:
 - a. contacting the sample with a nucleic acid probe that hybridizes under stringent hybridization conditions with genomic DNA from a plant comprising the DNA molecule of claim 1, and does not hybridize under such hybridization conditions with genomic DNA from an otherwise isogenic plant that does not comprise the recombinant nucleic acid molecule of claim 1, wherein the probe is homologous or complementary to SEQ ID NO:9, or a sequence that encodes a pesticidal protein comprising an amino acid sequence having at least 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO:10:
 - b. subjecting the sample and probe to stringent hybridization conditions; and
 - c, detecting hybridization of the probe with DNA of the sample.
- 26. A method of detecting the presence of a pesticidal protein or fragment thereof in a sample comprising protein, wherein said pesticidal protein comprises the amino acid sequence of SEQ ID NO: 10; or said pesticidal protein comprises an amino acid sequence having at least 90%, or 95%, or about 100% amino acid sequence identity to SEQ ID NO: 10; comprising:

83

- a. contacting the sample with an immunoreactive antibody; and
- b. detecting the presence of the protein.
- 27. The method of claim 26, wherein the step of detecting comprises an ELISA, or a Western blot.
- 28. The recombinant nucleic acid molecule of claim 1, wherein said pesticidal protein comprises the amino acid sequence of SEQ ID NO:10.
- **29**. The recombinant nucleic acid molecule of claim **1**, wherein said pesticidal protein comprises an amino acid 10 sequence having at least 90% amino acid sequence identity to SEQ ID NO:10.
- **30**. The recombinant nucleic acid molecule of claim **1**, wherein said pesticidal protein comprises an amino acid sequence having at least 95% amino acid sequence identity 15 to SEQ ID NO:10.
- **31**. The recombinant nucleic acid molecule of claim 1, wherein said pesticidal protein comprises an amino acid sequence having about 100% amino acid sequence identity to SEQ ID NO:10.
- **32**. The recombinant nucleic acid molecule of claim 1, wherein said polynucleotide segment hybridizes under stringent hybridization conditions to a polynucleotide having the nucleotide sequence of SEQ ID NO:9.

* * *